

	Type	Hits	Search Text	DBs	Time Stamp
1	BRS	15317	707/1-7,9,10.ccls.	US-PGPUB; USPAT	2005/03/24 09:24
2	BRS	10251	707/100-104.1.ccls.	US-PGPUB; USPAT	2005/03/24 09:25
3	BRS	2871	707/200,203.ccls.	US-PGPUB; USPAT	2005/03/24 09:25
4	BRS	9980	705/1,14,26,27.ccls.	US-PGPUB; USPAT	2005/03/24 09:26
5	BRS	22939	709/200,201-203,217-219,223- 229.ccls.	US-PGPUB; USPAT	2005/03/24 09:27
6	BRS	8849	715/500,501.1,513- 517,526,700,733,738,760,764,810 ,848,853,854,866,965-968.ccls.	US-PGPUB; USPAT	2005/03/24 09:32
7	BRS	5800	713/200,201.ccls.	US-PGPUB; USPAT	2005/03/24 09:32
8	BRS	1563	345/418,473.ccls.	US-PGPUB; USPAT	2005/03/24 09:33
9	BRS	23558	S1 or S2 or S3	US-PGPUB; USPAT	2005/03/30 08:17
10	BRS	46223	S4 or S5 or S7 or S6 or S8	US-PGPUB; USPAT	2005/03/30 08:17
11	BRS	15317	707/1-7,9,10.ccls.	US-PGPUB; USPAT	2005/03/25 08:18
12	BRS	10251	707/100-104.1.ccls.	US-PGPUB; USPAT	2005/03/25 08:18
13	BRS	2871	707/200,203.ccls.	US-PGPUB; USPAT	2005/03/25 08:18
14	BRS	23558	S11 or S12 or S13	US-PGPUB; USPAT	2005/03/25 08:18
15	BRS	9980	705/1,14,26,27.ccls.	US-PGPUB; USPAT	2005/03/25 08:18
16	BRS	22939	709/200,201-203,217-219,223- 229.ccls.	US-PGPUB; USPAT	2005/03/25 08:18
17	BRS	8849	715/500,501.1,513- 517,526,700,733,738,760,764,810 ,848,853,854,866,965-968.ccls.	US-PGPUB; USPAT	2005/03/25 08:18
18	BRS	5800	713/200,201.ccls.	US-PGPUB; USPAT	2005/03/25 08:18
19	BRS	1563	345/418,473.ccls.	US-PGPUB; USPAT	2005/03/25 08:18
20	BRS	46223	S15 or S16 or S18 or S17 or S19	US-PGPUB; USPAT	2005/03/25 08:18
21	BRS	2327	(accept\$4 with (query or queries)) and input\$4	US-PGPUB; USPAT	2005/03/25 08:20
22	BRS	33312	search\$4 and (data adj structure\$2)	US-PGPUB; USPAT	2005/03/25 08:21

23	BRS	1069391	advertiser\$2 and (web with page\$2) abd information	US-PGPUB; USPAT	2005/03/25 08:22
----	-----	---------	---	--------------------	---------------------

	Type	Hits	Search Text	DBs	Time Stamp
24	BRS	3087	advertiser\$2 and (web with page\$2) and information	US-PGPUB; USPAT	2005/03/25 08:22
25	BRS	8834	accept\$4 and (search\$4 with results) and generat\$4	US-PGPUB; USPAT	2005/03/25 08:23
26	BRS	1041	retriev\$4 with advertisement\$2	US-PGPUB; USPAT	2005/03/25 08:24
27	BRS	583	inverted adj2 (file\$2 or index\$2 or indices)	US-PGPUB; USPAT	2005/03/25 08:25
28	BRS	30606	world adj wide adj web	US-PGPUB; USPAT	2005/03/25 08:26
29	BRS	308	S21 and S22 and S25	US-PGPUB; USPAT	2005/03/25 08:27
30	BRS	1	S21 and S22 and S25 and S24 and S26	US-PGPUB; USPAT	2005/03/25 08:27
31	BRS	18	S21 and S22 and S25 and S24	US-PGPUB; USPAT	2005/03/25 08:28
32	BRS	18	S21 and S22 and S25 and S24 and retriev\$4	US-PGPUB; USPAT	2005/03/25 08:28
33	BRS	15	S21 and S22 and S25 and S24 and retriev\$4 and S28	US-PGPUB; USPAT	2005/03/29 13:40
34	BRS	2330	(accept\$4 with (query or queries)) and input\$4	US-PGPUB; USPAT	2005/03/29 13:40
35	BRS	33347	search\$4 and (data adj structure\$2)	US-PGPUB; USPAT	2005/03/29 13:40
36	BRS	8851	accept\$4 and (search\$4 with results) and generat\$4	US-PGPUB; USPAT	2005/03/29 13:40
37	BRS	30648	world adj wide adj web	US-PGPUB; USPAT	2005/03/29 13:40
38	BRS	136	S34 and S35 and S36 and retriev\$4 and S37	US-PGPUB; USPAT	2005/03/29 13:41
39	BRS	3	S34 and S35 and S36 and retriev\$4 and S37 and (inverted with (index\$2 or indices)) and (term\$2 with count\$4)	US-PGPUB; USPAT	2005/03/29 13:42
40	BRS	14	S34 and S35 and S36 and retriev\$4 and S37 and (inverted with (index\$2 or indices))	US-PGPUB; USPAT	2005/03/29 13:54
41	BRS	364	((ad\$2 or advertisement\$2) with perform\$6) and ((ad\$2 or advertisement\$2) with price\$2)	US-PGPUB; USPAT	2005/03/29 13:56
42	BRS	10	S34 and S36 and ((ad\$2 or advertisement\$2) with perform\$6) and ((ad\$2 or advertisement\$2) with price\$2)	US-PGPUB; USPAT	2005/03/29 13:56
43	BRS	15346	707/1-7,9,10.ccls.	US-PGPUB; USPAT	2005/03/30 08:17

44	BRS	10278	707/100-104.1.ccls.	US-PGPUB; USPAT	2005/03/30 08:17
----	-----	-------	---------------------	--------------------	---------------------

	Type	Hits	Search Text	DBs	Time Stamp
45	BRS	2879	707/200,203.ccls.	US-PGPUB; USPAT	2005/03/30 08:17
46	BRS	23605	S43 or S44 or S45	US-PGPUB; USPAT	2005/03/30 08:17
47	BRS	9995	705/1,14,26,27.ccls.	US-PGPUB; USPAT	2005/03/30 08:17
48	BRS	22982	709/200,201-203,217-219,223- 229.ccls.	US-PGPUB; USPAT	2005/03/30 08:17
49	BRS	8864	715/500,501.1,513- 517,526,700,733,738,760,764,810 ,848,853,854,866,965-968.ccls.	US-PGPUB; USPAT	2005/03/30 08:17
50	BRS	5812	713/200,201.ccls.	US-PGPUB; USPAT	2005/03/30 08:17
51	BRS	1567	345/418,473.ccls.	US-PGPUB; USPAT	2005/03/30 08:17
52	BRS	46302	S47 or S48 or S50 or S49 or S51	US-PGPUB; USPAT	2005/03/30 08:17

?ds,

Set	Items	Description
S1	73301	ADVERTIS? OR ADVERTIZ? OR AD OR ADS OR PROMOTION? OR ADVERT? T? ? OR COMMERCIAL()MESSAGE?
S2	1181	(INVERT?) (2N)(INDEX? OR INDICES OR FILE OR FILES OR FILI- NG? OR LIST OR LISTS OR LISTING? OR STRUCTURE?)
S3	21055	(INTERNET OR WEB OR ONLINE OR ON()LINE OR HOME) (2N) (PAGE OR PAGES OR SITE OR SITES OR PORTAL? OR DIRECTOR?)
S4	34902	(E OR ELECTRONIC OR DIGITAL OR VIRTUAL) (1W) (MAIL??? OR M- ESSAG??? OR CORRESPOND?) OR EMAIL???? OR (INTERNET OR ON()LI- NE OR ONLINE OR WEB) (1W) MAIL???? OR MIME OR SMTP OR POP(1N)- MAIL
S5	3	S1 AND S2
S6	1	S2 AND (S3 OR S4)
S7	17	S2 AND (INTERNET OR WEB OR WWW OR ONLINE OR ON()LINE)
S8	7	S7 AND IC=G06F?

?show files

File 347:JAPIO Nov 1976-2004/Aug(Updated 041203)

(c) 2004 JPO & JAPIO

File 350:Derwent WPIX 1963-2005/UD,UM &UP=200501

(c) 2005 Thomson Derwent

?

?ds

Set	Items	Description
S1	73140	ADVERTIS? OR ADVERTIZ? OR AD OR ADS OR PROMOTION? OR ADVERT- T? ? OR COMMERCIAL()MESSAGE?
S2	1938	(INVERT? OR FORWARD) (2N) (INDEX? ? OR INDICES OR FILE OR F- ILES OR FILING? OR LIST OR LISTS OR LISTING? OR STRUCTURE?)
S3	20991	(INTERNET OR WEB OR ONLINE OR ON()LINE OR HOME) (2N) (PAGE OR PAGES OR SITE OR SITES OR PORTAL? OR DIRECTOR?)
S4	29975	(E OR ELECTRONIC) (1W) (MAIL??? OR MESSAG??? OR CORRESPOND- ?) OR EMAIL???? OR (INTERNET OR ON()LINE OR ONLINE OR WEB) (- 1W) MAIL????
S5	5	S1 AND S2
S6	10	S3:S4 AND S2
S7	42	S2 AND (WEB OR INTERNET OR WWW OR ONLINE OR ON()LINE)
S8	1	S7 AND S1
S9	1167	INVERT?(2N) (INDEX? ? OR INDICES OR FILE OR FILES OR FILIN- G? OR LIST OR LISTS OR LISTING? OR STRUCTURE?)
S10	17	S9 AND (WEB OR INTERNET OR WWW OR ONLINE OR ON()LINE)
S11	7	S10 AND IC=G06F?

?show files

File 347:JAPIO Nov 1976-2004/Aug(Updated 041203)

(c) 2004 JPO & JAPIO

File 350:Derwent WPIX 1963-2004/UD,UM &UP=200482

(c) 2004 Thomson Derwent

?

?t s6/5/7

6/5/7 (Item 7 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2004 Thomson Derwent. All rts. reserv.

014004914 **Image available**
WPI Acc No: 2001-489128/200154
XRPX Acc No: N01-361923

Method for storing information about web documents such as pages or sites
in a manner that may be used in conjunction with inverted term lists
to facilitate the retrieval of documents of interest from the web
Patent Assignee: GTE LAB INC (SYLV); VERIZON LAB INC (VERI-N)

Inventor: PONTE J M

Number of Countries: 002 Number of Patents: 002

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
CA 2310931	A1	20010130	CA 2310931	A	20000607	200154 B
US 6665665	B1	20031216	US 99365326	A	19990730	200382

Priority Applications (No Type Date): US 99365326 A 19990730

Patent Details:

Patent No	Kind	Lan Pg	Main IPC	Filing Notes
CA 2310931	A1	E 105	G06F-017/30	
US 6665665	B1		G06F-017/30	

Abstract (Basic): CA 2310931 A1

NOVELTY - The method for storing information about Web documents such as pages or sites in a manner that may be used in conjunction with **inverted term lists** to facilitate the retrieval of documents of interest from the Web. The method involves constructing compressed surrogates for each document in the database and inserting in the compressed documents surrogate information about terms which occur in the document, such that various operations may be performed without the need to retrieve a copy of the document from the Web. **Inverted term lists** that contain information about terms that occur in the database are also created in conjunction with creation of the compressed document surrogates.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for: (i) a device for maintaining information about a collection of documents in a data base to facilitate determining which documents may be of interest; (ii) a device for modifying a collection of **inverted term lists**; (iii) a device for determining the score for a document under a search query which specifies terms that are desired to be present or absent; (iv) a device for returning a list of a desired number of documents N in order of predicted utility, from among a collection of documents, as predicted by a search query containing terms desired to be present or absent.

USE - For maintaining information about material on the World Wide Web to facilitate retrieval of **web pages** of interest to a user that relate to electronic commerce.

ADVANTAGE - Permits efficient updating of **inverted term lists** when documents on the Web have been modified or deleted, and also permits the efficient processing of search queries in a variety of circumstances.

DESCRIPTION OF DRAWING(S) - The drawing shows a schematic diagram of the computer system.

computer system (1)
pp; 105 DwgNo 1/14

Title Terms: METHOD; STORAGE; INFORMATION; WEB; DOCUMENT; PAGE; SITE;
MANNER; CONJUNCTION; INVERT; TERM; LIST; FACILITATE; RETRIEVAL; DOCUMENT;
INTEREST; WEB

Derwent Class: T01; W01

International Patent Class (Main): G06F-017/30

International Patent Class (Additional): G06F-007/00; H04L-012/24

File Segment: EPI

?

?t s11/5/1-7

11/5/1 (Item 1 from file: 347)
DIALOG(R)File 347:JAPIO
(c) 2004 JPO & JAPIO. All rts. reserv.

06103734 **Image available**

WEB DOCUMENT RETRIEVAL SUPPORTING DEVICE AND COMPUTER READABLE RECORDING
MEDIUM RECORDED WITH PROGRAM FOR FUNCTIONING COMPUTER AS THE DEVICE

PUB. NO.: 11-045257 [JP 11045257 A]
PUBLISHED: February 16, 1999 (19990216)
INVENTOR(s): WAKASUGI TAKASHI
APPLICANT(s): JUST SYST CORP
APPL. NO.: 09-199618 [JP 97199618]
FILED: July 25, 1997 (19970725)
INTL CLASS: G06F-017/30

ABSTRACT

PROBLEM TO BE SOLVED: To automatically classify gathered **Web** documents into prepared respective categories and to reduce labor required for the classifying work of the **Web** documents.

SOLUTION: This device is provided with a **Web** document gathering software 104 for gathering the **Web** documents, an **inverted file** 109 for storing retrieval information used for retrieving the gathered **Web** documents, a category management software 105 for inputting retrieval conditions, setting the inputted retrieval conditions as classification items and presenting the classification items corresponding to a request from a **Web** client and a retrieval software 106 for retrieving the pertinent **Web** document by using the retrieval information stored in the **inverted file** 109 based on the selected classification item when the classification item is selected in the **Web** client. In this case, the category management software 105 presents the list of the pertinent **Web** documents to the **Web** client based on the retrieved result of the retrieval software 106.

COPYRIGHT: (C)1999, JPO

11/5/2 (Item 2 from file: 347)
DIALOG(R)File 347:JAPIO
(c) 2004 JPO & JAPIO. All rts. reserv.

03060071 **Image available**

ON - LINE INFORMATION RETRIEVAL SUPPORTING DEVICE

PUB. NO.: 02-035571 [JP 2035571 A]
PUBLISHED: February 06, 1990 (19900206)
INVENTOR(s): MORITA TETSUYA
APPLICANT(s): RICOH CO LTD [000674] (A Japanese Company or Corporation), JP
(Japan)
APPL. NO.: 63-184477 [JP 88184477]
FILED: July 26, 1988 (19880726)
INTL CLASS: [5] G06F-015/40 ; G06F-012/00
JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications); 45.2
(INFORMATION PROCESSING -- Memory Units)
JOURNAL: Section: P, Section No. 1037, Vol. 14, No. 189, Pg. 161,
April 17, 1990 (19900417)

ABSTRACT

PURPOSE: To efficiently execute high-speed retrieval by down loading information from an **on - line** data base after limiting a retrieval object file to proper quantity by a private file.

CONSTITUTION: For the side of the terminal of a user, a private data base 20 having an inherent keyword file 17 peculiar to the user, a covalent keyword file 18, and an inherent **inverted file** 19 to indicate the corresponding relations between an inherent keyword and the **on - line**

data base is provided. When the on - line data base is to be retrieved, first, the retrieval object file is limited to the proper quantity by the private data base 20 with a command and a data format standardized for the user, and after that, the information is down loaded from the on - line data base with the standardized command for the user.

11/5/3 (Item 1 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2004 Thomson Derwent. All rts. reserv.

016606405 **Image available**
WPI Acc No: 2004-765139/200475
XRPX Acc No: N04-603601

Inverted index storing method for internet searching, involves storing index information related to same index item in continuous blocks and using index units in each index block for storing index information related to same index item

Patent Assignee: INT BUSINESS MACHINES CORP (IBMC)

Inventor: PAN Y; SU Z; YANG L P

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
US 20040205044	A1	20041014	US 2004818833	A	20040406	200475 B

Priority Applications (No Type Date): CN 2003109847 A 20030411

Patent Details:

Patent No	Kind	Lan Pg	Main IPC	Filing Notes
US 20040205044	A1	18	G06F-017/30	

US 20040205044 A1 18 G06F-017/30

Abstract (Basic): US 20040205044 A1

NOVELTY - The index information related to each index item is sequentially stored in the newly created **inverted file**, such that the index information related to the same index item is stored in continuous blocks, and the index units in each index block are only used for storing the index information related to the same index item.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

- (1) program product for storing **inverted index**; and
- (2) **inverted index** mechanism.

USE - For fill-text retrieval in **internet** search.

ADVANTAGE - No need to relocate the reading pointer to the file, when reading the index information on an arbitrarily chosen index item, thus reducing the file reading time. When performing an operation on the index information in an index block, other index items are not affected, thus it is possible to **on - line** update the index information in any index block through a simple locking-unlocking method, without having to stop searching service.

DESCRIPTION OF DRAWING(S) - The figure shows flowcharts of method for storing **inverted index**.
pp; 18 DwgNo 1C/8

Title Terms: INVERT; INDEX; STORAGE; METHOD; SEARCH; STORAGE; INDEX; INFORMATION; RELATED; INDEX; ITEM; CONTINUOUS; BLOCK; INDEX; UNIT; INDEX; BLOCK; STORAGE; INDEX; INFORMATION; RELATED; INDEX; ITEM

Derwent Class: T01

International Patent Class (Main): G06F-017/30

File Segment: EPI

11/5/4 (Item 2 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2004 Thomson Derwent. All rts. reserv.

015460527 **Image available**
WPI Acc No: 2003-522669/200349
XRPX Acc No: N03-414759

Internet based data record search method in database, involves constructing query corresponding to given search criteria and executing

it on identified regions of database

Patent Assignee: MICROSOFT CORP (MICT)

Inventor: AGRAWAL S; CHAUDHURI S

Number of Countries: 001 Number of Patents: 002

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
US 20030078915	A1	20030424	US 200136348	A	20011019	200349 B
US 6792414	B2	20040914	US 200136348	A	20011019	200460

Priority Applications (No Type Date): US 200136348 A 20011019

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
US 20030078915	A1		26	G06F-007/00	
US 6792414	B2			G06F-017/30	

Abstract (Basic): US 20030078915 A1

NOVELTY - An **inverter list** of keywords that maps the data record components to a region of database containing corresponding data record, is created. The regions of database containing data records relating to the given search keyword, are identified by accessing the **inverted list**. A query is constructed corresponding to the given search criteria and is executed on the identified regions of database.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

(1) computer readable medium storing instructions to perform data record search process; and

(2) data records search apparatus.

USE - For searching data records in database comprising address information of employee, mailing list information, product and sales details.

ADVANTAGE - The records matching the search criteria are efficiently retrieved by executing the query on the identified regions. The keyword searching on relational database is made efficient.

DESCRIPTION OF DRAWING(S) - The figure shows the flowchart of data record search process.

pp; 26 DwgNo 6/24

Title Terms: BASED; DATA; RECORD; SEARCH; METHOD; DATABASE; CONSTRUCTION;

QUERY; CORRESPOND; SEARCH; CRITERIA; EXECUTE; IDENTIFY; REGION; DATABASE

Derwent Class: T01

International Patent Class (Main): G06F-007/00 ; G06F-017/30

File Segment: EPI

11/5/5 (Item 3 from file: 350)

DIALOG(R)File 350:Derwent WPIX

(c) 2004 Thomson Derwent. All rts. reserv.

015369032 **Image available**

WPI Acc No: 2003-429970/200340

XRPX Acc No: N03-343391

Interactive multimedia delivery system used in networking system, has retriever that exploits specific statistics maintained by inverted indices , to rank relevance of nodes for new annotation set

Patent Assignee: KNUMI INC (KNUM-N)

Inventor: DEY J K; SIVASANKARAN R M

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
US 20030061028	A1	20030327	US 2001956889	A	20010921	200340 B

Priority Applications (No Type Date): US 2001956889 A 20010921

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
US 20030061028	A1		17	G06F-017/27	

Abstract (Basic): US 20030061028 A1

NOVELTY - The system derives data from the previous mappings of annotations to nodes in an ontology. A retriever exploits specific

statistics maintained by **inverted indices** , to rank the relevance of the nodes for a new annotation set. The information related to new annotation set. The information related to new annotation, is extracted from the most relevant node and database through network.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

- (1) ontology searching method;
- (2) contextual information retrieval method;
- (3) contextual information retrieval system;
- (4) annotation mapping method; and
- (5) article of manufacture comprises computer medium with computer readable program code for searching ontology.

USE - Interactive multimedia delivery system implemented on multi-nodal system e.g. LAN or networking system e.g. **Internet** , world wide **web** (**WWW**), wireless **web** , for delivering annotations through television, computer handheld device or telephone.

ADVANTAGE - The multimedia authoring environment enables broadband producer to rapidly create a document that integrates multimedia content with other content that is relevant to the multimedia segment.

DESCRIPTION OF DRAWING(S) - The figure shows the ways for obtaining various multimedia document annotations.

pp; 17 DwgNo 8/8

Title Terms: INTERACT; DELIVER; SYSTEM; SYSTEM; RETRIEVAL; EXPLOIT;

SPECIFIC; STATISTICAL; MAINTAIN; INVERT; INDEX; RANK; RELEVANT; NODE; NEW ; SET

Derwent Class: T01; W02

International Patent Class (Main): G06F-017/27

File Segment: EPI

11/5/6 (Item 4 from file: 350)

DIALOG(R)File 350:Derwent WPIX

(c) 2004 Thomson Derwent. All rts. reserv.

014004914 **Image available**

WPI Acc No: 2001-489128/200154

XRPX Acc No: N01-361923

Method for storing information about web documents such as pages or sites in a manner that may be used in conjunction with inverted term lists to facilitate the retrieval of documents of interest from the web

Patent Assignee: GTE LAB INC (SYLV); VERIZON LAB INC (VERI-N)

Inventor: PONTE J M

Number of Countries: 002 Number of Patents: 002

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
CA 2310931	A1	20010130	CA 2310931	A	20000607	200154 B
US 6665665	B1	20031216	US 99365326	A	19990730	200382

Priority Applications (No Type Date): US 99365326 A 19990730

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
-----------	------	-----	----	----------	--------------

CA 2310931	A1	E	105	G06F-017/30	
------------	----	---	-----	-------------	--

US 6665665	B1			G06F-017/30	
------------	----	--	--	-------------	--

Abstract (Basic): CA 2310931 A1

NOVELTY - The method for storing information about **Web** documents such as pages or sites in a manner that may be used in conjunction with **inverted term lists** to facilitate the retrieval of documents of interest from the **Web** . The method involves constructing compressed surrogates for each document in the database and inserting in the compressed documents surrogate information about terms which occur in the document, such that various operations may be performed without the need to retrieve a copy of the document from the **Web** . **Inverted term lists** that contain information about terms that occur in the database are also created in conjunction with creation of the compressed document surrogates.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for: (i) a device for maintaining information about a collection of documents in a

data base to facilitate determining which documents may be of interest;
(ii) a device for modifying a collection of **inverted term lists** ;
(iii) a device for determining the score for a document under a search query which specifies terms that are desired to be present or absent;
(iv) a device for returning a list of a desired number of documents N in order of predicted utility, from among a collection of documents, as predicted by a search query containing terms desired to be present or absent.

USE - For maintaining information about material on the World Wide Web to facilitate retrieval of **web** pages of interest to a user that relate to electronic commerce.

ADVANTAGE - Permits efficient updating of **inverted term lists** when documents on the **Web** have been modified or deleted, and also permits the efficient processing of search queries in a variety of circumstances.

DESCRIPTION OF DRAWING(S) - The drawing shows a schematic diagram of the computer system.

computer system (1)

pp; 105 DwgNo 1/14

Title Terms: METHOD; STORAGE; INFORMATION; **WEB** ; DOCUMENT; PAGE; SITE; MANNER; CONJUNCTION; INVERT; TERM; LIST; FACILITATE; RETRIEVAL; DOCUMENT; INTEREST; **WEB**

Derwent Class: T01; W01

International Patent Class (Main): G06F-017/30

International Patent Class (Additional): G06F-007/00 ; H04L-012/24

File Segment: EPI

11/5/7 (Item 5 from file: 350)

DIALOG(R)File 350:Derwent WPIX

(c) 2004 Thomson Derwent. All rts. reserv.

011618032 **Image available**

WPI Acc No: 1998-035160/199804

XRPX Acc No: N98-028228

Query evaluation system for very large full text database on Internet -
has data structure in static cache contg. data about each document to
calculate score for important documents for large

Patent Assignee: EXCITE INC (EXCI-N)

Inventor: SPENCER G

Number of Countries: 019 Number of Patents: 002

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
EP 813158	A2	19971217	EP 97109680	A	19970613	199804 B
US 5915249	A	19990622	US 96661335	A	19960614	199931

Priority Applications (No Type Date): US 96661335 A 19960614

Patent Details:

Patent No	Kind	Lan Pg	Main IPC	Filing Notes
-----------	------	--------	----------	--------------

EP 813158	A2	E 26	G06F-017/30	
-----------	----	------	-------------	--

Designated States (Regional): AT BE CH DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE

US 5915249	A	G06F-017/30
------------	---	-------------

Abstract (Basic): EP 813158 A

The database access system provides query searches into very large full-text document databases, e.g. **Internet** search agents. The search system uses an **inverted index** database containing document and term frequency data. A static cache is created from the **inverted index**. This contains entries in contribution order for **inverted index** entries with large numbers of associated documents.

Each term of a query is initially searched in the static cache and term contributions combined. An additional look-up table is included in the cache to aid in accessing the **inverted index** database.

ADVANTAGE - Provides more rapid access and query evaluation of search database than directly evaluating **inverted index**.

Dwg.1/5

Title Terms: QUERY; EVALUATE; SYSTEM; FULL; TEXT; DATABASE; DATA; STRUCTURE

; STATIC; CACHE; CONTAIN; DATA; DOCUMENT; CALCULATE; SCORE; IMPORTANT;
DOCUMENT

Derwent Class: T01

International Patent Class (Main): G06F-017/30

File Segment: EPI

?

?ds

Set	Items	Description
S1	786112	ADVERTIS? OR ADVERTIZ? OR AD OR ADS OR PROMOTION? OR ADVERT- T? ? OR COMMERCIAL()MESSAGE?
S2	8273	(INVERT?) (2N)(INDEX? OR INDICES OR FILE OR FILES OR FILI- NG? OR LIST OR LISTS OR LISTING? OR STRUCTURE?)
S3	45	S1 AND S2
S4	4	S3 AND (INTERNET OR WEB????? OR WWW OR ONLINE OR ON()LINE)
S5	2	RD (unique items)
S6	140817	(INTERNET OR WEB OR ONLINE OR ON()LINE OR HOME) (2N) (PAGE OR PAGES OR SITE OR SITES OR PORTAL? OR DIRECTOR?)
S7	17	S2 AND S6
S8	11	RD (unique items)

?show file

File 8: Ei Compendex(R) 1970-2005/Dec W4
(c) 2005 Elsevier Eng. Info. Inc.

File 35: Dissertation Abs Online 1861-2004/Dec
(c) 2004 ProQuest Info&Learning

File 103: Energy SciTec 1974-2004/Dec B2
(c) 2005 Contains copyrighted material

File 65: Inside Conferences 1993-2005/Jan W1
(c) 2005 BLDSC all rts. reserv.

File 2: INSPEC 1969-2004/Dec W2
(c) 2004 Institution of Electrical Engineers

File 94: JICST-EPlus 1985-2004/Nov W4
(c) 2004 Japan Science and Tech Corp(JST)

File 438: Library Lit. & Info. Science 1984-2004/Oct
(c) 2004 The HW Wilson Co

File 111: TGG Natl. Newspaper Index(SM) 1979-2005/Jan 04
(c) 2005 The Gale Group

File 603: Newspaper Abstracts 1984-1988
(c) 2001 ProQuest Info&Learning

File 483: Newspaper Abs Daily 1986-2005/Jan 05
(c) 2005 ProQuest Info&Learning

File 6: NTIS 1964-2004/Dec W4
(c) 2004 NTIS, Intl Cpyrght All Rights Res

File 144: Pascal 1973-2004/Dec W1
(c) 2004 INIST/CNRS

File 434: SciSearch(R) Cited Ref Sci 1974-1989/Dec
(c) 1998 Inst for Sci Info

File 34: SciSearch(R) Cited Ref Sci 1990-2005/Jan W1
(c) 2005 Inst for Sci Info

File 62: SPIN(R) 1975-2004/Oct W4
(c) 2004 American Institute of Physics

File 99: Wilson Appl. Sci & Tech Abs 1983-2004/Nov
(c) 2004 The HW Wilson Co.

File 266: FEDRIP 2004/Sep
Comp & dist by NTIS, Intl Copyright All Rights Res

File 474: New York Times Abs 1969-2005/Jan 05
(c) 2005 The New York Times

File 475: Wall Street Journal Abs 1973-2005/Jan 04
(c) 2005 The New York Times

?

?t.s8/3,k/5,7,8

8/3,K/5 (Item 3 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2004 Institution of Electrical Engineers. All rts. reserv.

7236626 INSPEC Abstract Number: C2002-05-7250L-003

Title: Building a distributed full-text index for the Web

Author(s): Melnik, S.; Raghavan, S.; Yang, B.; Garcia-Molina, H.

Author Affiliation: Dept. of Comput. Sci., Stanford Univ., CA, USA

Journal: ACM Transactions on Information Systems vol.19, no.3 p. 217-41

Publisher: ACM,

Publication Date: July 2001 Country of Publication: USA

CODEN: ATISET ISSN: 1046-8188

SICI: 1046-8188(200107)19:3L:217:BDFT;1-G

Material Identity Number: N617-2002-001

U.S. Copyright Clearance Center Code: 1046-8188/07/0100-0217\$5.00

Language: English

Subfile: C

Copyright 2002, IEE

Abstract: We identify crucial design issues in building a distributed **inverted index** for a large collection of **Web pages**. We introduce a novel pipelining technique for structuring the core index-building system that substantially reduces the index construction time. We also propose a storage scheme for creating and managing **inverted files** using an embedded database system. We suggest and compare different strategies for collecting global statistics from distributed **inverted indexes**. Finally, we present performance results from experiments on a testbed distributed Web indexing system that...

Identifiers: distributed **inverted index** ; ...

... **Web pages** ; ...

... **inverted files** ;

8/3,K/7 (Item 5 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2004 Institution of Electrical Engineers. All rts. reserv.

5583971 INSPEC Abstract Number: C9706-7250R-016

Title: Smart systems, smart searches

Author(s): Brunelle, B.S.

Author Affiliation: Ovid Technols., New York, NY, USA

Conference Title: Online Information 96 Proceedings. 20th International Online Information Meeting p.387-90

Editor(s): Raitt, D.I.; Jeapes, B.

Publisher: Learned Inf. (Europe), Oxford, UK

Publication Date: 1996 Country of Publication: UK xviii+487 pp.

ISBN: 1 900871 04 1 Material Identity Number: XX96-03405

Conference Title: Proceedings of Online Information 96

Conference Sponsor: The Guardian

Conference Date: 3-5 Dec. 1996 Conference Location: London, UK

Language: English

Subfile: C

Copyright 1997, IEE

...Abstract: R solutions being offered on the Web are quite familiar to the librarian. The various **Web search sites** make use of "traditional" **inverted indexes**, manual **indexing**, automatic indexing based on statistical models, relevance ranking and document clustering. All of these statistical...

...Identifiers: **Web search sites** ; ...

... **inverted indexes** ;

8/3,K/8 (Item 1 from file: 144)
DIALOG(R)File 144:Pascal
(c) 2004 INIST/CNRS. All rts. reserv.

16899475 PASCAL No.: 04-0561910
Index construction for linear categorisation
Twelfth ACM international conference on information & knowledge
management : New Orleans LA, 3-8 November 2003
SHANKS Vaughan R; WILLIAMS Hugh E
FRIEDER Ophir, ed; HAMMER Joachim, ed; QHERSHI Sajda, ed; SELIGMAN Len,
ed
School of Computer Science and Information Technology, RMIT University,
GPO Box 2476V, Melbourne 3001, Australia
ACM. SIGIR special interest group on information retrieval, United States
; ACM. SIGMIS special interest group on management information systems,
United States
CIKM 2003 : conference on information and knowledge management, 12 (New
Orleans LA USA) 2003-11-03
2003 334-341
Publisher: Association for Computing Machinery, New York NY
Language: English

Copyright (c) 2004 INIST-CNRS. All rights reserved.

... can be browsed or searched to more accurately and quickly meet
information needs. On the **Web**, category-based **portals** such as Yahoo!
and DMOZ are extremely popular: DMOZ is maintained by over 56,000...

... collections such as DMOZ or large news wire feeds. In this paper, we
show how **inverted indexes** can be used for scalable training in
categorisation, and propose novel heuristics for a fast...

English Descriptors: **Index** (documentation); **Inverted file** ;
Construction; Performance evaluation
?

?ds

Set	Items	Description
S1	684246	ADVERTIS? OR ADVERTIZ? OR AD OR ADS OR PROMOTION? OR ADVERT? T? ? OR COMMERCIAL()MESSAGE?
S2	8200	(INVERT?) (2N) (INDEX? ? OR INDICES OR FILE OR FILES OR FI- LING? OR LIST OR LISTS OR LISTING? OR STRUCTURE?)
S3	131644	(INTERNET OR WEB OR ONLINE OR ON()LINE OR HOME) (2N) (PAGE OR PAGES OR SITE OR SITES OR PORTAL? OR DIRECTOR?)
S4	93781	(E OR ELECTRONIC OR DIGITAL OR VIRTUAL) (1W) (MAIL??? OR M- ESSAG??? OR CORRESPOND?) OR EMAIL???? OR (INTERNET OR ON()LI- NE OR ONLINE OR WEB) (1W) MAIL???? OR MIME OR SMTP OR POP(1N)- MAIL
S5	0	S1 AND S2 AND S3:S4
S6	19	S2 AND S3:S4
S7	14	RD (unique items)

?show files

File 8: Ei Compendex(R) 1970-2005/Dec W4
(c) 2005 Elsevier Eng. Info. Inc.

File 35: Dissertation Abs Online 1861-2004/Dec
(c) 2004 ProQuest Info&Learning

File 103: Energy SciTec 1974-2004/Dec B1
(c) 2004 Contains copyrighted material

File 65: Inside Conferences 1993-2005/Jan W1
(c) 2005 BLDSC all rts. reserv.

File 2: INSPEC 1969-2004/Dec W2
(c) 2004 Institution of Electrical Engineers

File 94: JICST-EPlus 1985-2004/Nov W4
(c) 2004 Japan Science and Tech Corp(JST)

File 438: Library Lit. & Info. Science 1984-2004/Oct
(c) 2004 The HW Wilson Co

File 111: TGG Natl. Newspaper Index(SM) 1979-2005/Jan 03
(c) 2005 The Gale Group

File 603: Newspaper Abstracts 1984-1988
(c) 2001 ProQuest Info&Learning

File 483: Newspaper Abs Daily 1986-2005/Jan 04
(c) 2005 ProQuest Info&Learning

File 6: NTIS 1964-2004/Dec W4
(c) 2004 NTIS, Intl Cpyrght All Rights Res

File 144: Pascal 1973-2004/Dec W1
(c) 2004 INIST/CNRS

File 434: SciSearch(R) Cited Ref Sci. 1974-1989/Dec
(c) 1998 Inst for Sci Info

File 34: SciSearch(R) Cited Ref Sci 1990-2005/Jan W1
(c) 2005 Inst for Sci Info

File 62: SPIN(R) 1975-2004/Oct W4
(c) 2004 American Institute of Physics

File 99: Wilson Appl. Sci & Tech Abs 1983-2004/Nov
(c) 2004 The HW Wilson Co.

File 266: FEDRIP 2004/Sep
Comp & dist by NTIS, Intl Copyright All Rights Res

?

?t s7/5/1-12,14

7/5/1 (Item 1 from file: 8)
DIALOG(R)File 8:Ei Compendex(R)
(c) 2005 Elsevier Eng. Info. Inc. All rts. reserv.

06460526 E.I. No: EIP03307561500

Title: I/O-efficient techniques for computing pagerank
Author: Chen, Yen-Yu; Gan, Qingqing; Suel, Torsten
Corporate Source: CIS Department Polytechnic University, Brooklyn, NY
11201, United States
Conference Title: Proceedings of the Eleventh International Conference on
Information and Knowledge Management (CIKM 2002)
Conference Location: McLean, VA, United States Conference Date:
20021104-20021109
Sponsor: ACM; SIGIR; SIGMIS
E.I. Conference No.: 61158
Source: International Conference on Information and Knowledge Management,
Proceedings 2002. p 549-557
Publication Year: 2002
Language: English
Document Type: CA; (Conference Article) Treatment: T; (Theoretical)
Journal Announcement: 0307W4

Abstract: Over the last few years, most major search engines have integrated link-based ranking techniques in order to provide more accurate search results. One widely known approach is the Pagerank technique, which forms the basis of the Google ranking scheme, and which assigns a global importance measure to each page based on the importance of other pages pointing to it. The main advantage of the Pagerank measure is that it is independent of the query posed by a user; this means that it can be precomputed and then used to optimize the layout of the **inverted index structure** accordingly. However, computing the Pagerank measure requires implementing an iterative process on a massive graph corresponding to billions of **web pages** and hyperlinks. In this paper, we study I/O-efficient techniques to perform this iterative computation. We derive two algorithms for Pagerank based on techniques proposed for out-of-core graph algorithms, and compare them to two existing algorithms proposed by Haveliwala. We also consider the implementation of a recently proposed topic-sensitive version of Pagerank. Our experimental results show that for very large data sets, significant improvements over previous results can be achieved on machines with moderate amounts of memory. On the other hand, at most minor improvements are possible on data sets that are only moderately larger than memory, which is the case in many practical scenarios. 43 Refs.

Descriptors: *World Wide Web; Search engines; Iterative methods;
Algorithms; Graph theory; Optimization

Identifiers: Pagerank technique; External memory algorithms

Classification Codes:

723.5 (Computer Applications); 921.6 (Numerical Methods); 921.4
(Combinatorial Mathematics, Includes Graph Theory, Set Theory); 921.5
(Optimization Techniques)

723 (Computer Software, Data Handling & Applications); 921 (Applied
Mathematics)

72 (COMPUTERS & DATA PROCESSING); 92 (ENGINEERING MATHEMATICS)

7/5/2 (Item 1 from file: 35)
DIALOG(R)File 35:Dissertation Abs Online
(c) 2004 ProQuest Info&Learning. All rts. reserv.

01947193 ORDER NO: AADAA-IMQ77916

An **inverted index generator for Cindi**

Author: Li, Hudong

Degree: M.Comp.Sc.

Year: 2003

Corporate Source/Institution: Concordia University (Canada) (0228)

Adviser: B. C. Desai

Source: VOLUME 41/06 of MASTERS ABSTRACTS.

Descriptors: COMPUTER SCIENCE

Descriptor Codes: 0984

ISBN: 0-612-77916-5

Human maintained search engines are expensive, slow to update, and cannot cover all the **web pages**. Automated search engines that rely on keyword matching usually return too many low quality results, with most users only looking at the first few tens of the search results. Because search engine development has gone on at companies with little publication of technical details, it is a challenging task to develop a search engine. The use of hypertextual information can help to improve search quality. This report addresses the question of how to build an **inverted index** for a search system that can use the additional information presented in hypertext to produce better search results. This report is part of the work of the Concordia INDEXing and DIScovery (CINDI) Digital Library System. In this report, we summarize the research work I have done; we present some implementation issues for the project; and present the data structures that can be used in indexing **web pages**. The design decision was driven by the desire to have a reasonable compact data structure, and the ability to fetch a record in few disk seeks during a search. This project has been implemented in C++ on Linux platform.

7/5/3 (Item 1 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2004 Institution of Electrical Engineers. All rts. reserv.

8190805 INSPEC Abstract Number: C2005-01-7250N-010

Title: The comparison of modern search engines

Author(s): Owczarek, M.; Sakowicz, B.; Napieralski, A.

Conference Title: Modern Problems of Radio Engineering, Telecommunications and Computer Science. Proceedings of the International Conference TCSET 2004 (IEEE Cat. No.04EX822) p.333-6

Publisher: Lviv Politytechnic, Lviv, Ukraine

Publication Date: 2004 Country of Publication: Ukraine 632 pp.

ISBN: 966 553 380 0 Material Identity Number: XX-2004-02217

Conference Title: Modern Problems of Radio Engineering, Telecommunications and Computer Science. Proceedings of the International Conference TCSET 2004

Conference Sponsor: VAT Urktelecom

Conference Date: 24-28 Feb. 2004 Conference Location: Lviv-Slavsko, Ukraine

Language: English Document Type: Conference Paper (PA)

Treatment: Practical (P); Theoretical (T)

Abstract: At present time most explorers are created as simple **inverted index** and system ranking, which gives the largest speed and efficiency of explorer. This article is the comparison and classification of modern systems for indexing **internet sites**, classification, building and proposed solutions. (3 Refs)

Subfile: C

Descriptors: indexing; Internet; search engines

Identifiers: modern search engines; **Internet sites**; search explorer; system ranking; indexing

Class Codes: C7250N (Search engines); C7240 (Information analysis and indexing); C7210N (Information networks)

Copyright 2004, IEE

7/5/4 (Item 2 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2004 Institution of Electrical Engineers. All rts. reserv.

7274381 INSPEC Abstract Number: C2002-06-7250R-041

Title: Low-cost parallel text retrieval using PC-cluster

Author(s): Rungsawang, A.; Laohakanniyom, A.; Lertprasertkune, M.

Author Affiliation: Massive Inf. & Knowledge Eng., Kasetsart Univ., Bangkok, Thailand

Conference Title: Recent Advances in Parallel Virtual Machine and Message Passing Interface. 8th European PVM/MPI Users' Group Meeting. Proceedings (Lecture Notes in Computer Science Vol.2131) p.419-26

Editor(s): Cotronis, Y.; Dongarra, J.

Publisher: Springer-Verlag, Berlin, Germany

Publication Date: 2001 Country of Publication: Germany xv+438 pp.

ISBN: 3 540 42609 4 Material Identity Number: XX-2001-02418

Conference Title: Recent Advances in Parallel Virtual Machine and Message Passing Interface. 8th European PVM/MPI Users' Group Meeting. Proceedings

Conference Date: 23-26 Sept. 2001 Conference Location: Santorini/Thera, Greece

Language: English Document Type: Conference Paper (PA)

Treatment: Practical (P); Experimental (X)

Abstract: We present a parallel vector space based text retrieval prototype implemented on a low-cost PC cluster running Linux operating system, using the PVM message passing library. We also embed the **inverted file structure** into our proposed prototype for fast retrieval. From several experiments derived from the standard TREC-9 collection, this prototype can index up to 500000 **Web pages** per hour using a simple x86 machine. We also obtain 5.4 seconds query response time on searching in the one and a half million TREC-9 **Web pages**, using two machines. (15 Refs)

Subfile: C

Descriptors: indexing; information retrieval; message passing; parallel programming; text analysis; virtual machines; workstation clusters

Identifiers: low-cost parallel text retrieval; PC-cluster; Linux; PVM message passing library; **inverted file structure**; TREC-9 collection; **web pages**; query response time

Class Codes: C7250R (Information retrieval techniques); C6130D (Document processing techniques); C6110P (Parallel programming); C7430 (Computer engineering); C6150N (Distributed systems software); C6150J (Operating systems); C5620L (Local area networks); C7240 (Information analysis and indexing)

Copyright 2002, IEE

7/5/5 (Item 3 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2004 Institution of Electrical Engineers. All rts. reserv.

7236626 INSPEC Abstract Number: C2002-05-7250L-003

Title: Building a distributed full-text index for the Web

Author(s): Melnik, S.; Raghavan, S.; Yang, B.; Garcia-Molina, H.

Author Affiliation: Dept. of Comput. Sci., Stanford Univ., CA, USA

Journal: ACM Transactions on Information Systems vol.19, no.3 p. 217-41

Publisher: ACM,

Publication Date: July 2001 Country of Publication: USA

CODEN: ATISET ISSN: 1046-8188

SICI: 1046-8188(200107)19:3L:217:BDFT;1-G

Material Identity Number: N617-2002-001

U.S. Copyright Clearance Center Code: 1046-8188/07/0100-0217\$5.00

Language: English Document Type: Journal Paper (JP)

Treatment: Practical (P)

Abstract: We identify crucial design issues in building a distributed **inverted index** for a large collection of **Web pages**. We introduce a novel pipelining technique for structuring the core index-building system that substantially reduces the index construction time. We also propose a storage scheme for creating and managing **inverted files** using an embedded database system. We suggest and compare different strategies for collecting global statistics from distributed inverted indexes. Finally, we present performance results from experiments on a testbed distributed Web indexing system that we have implemented. (36 Refs)

Subfile: C

Descriptors: database indexing; full-text databases; Internet; parallel algorithms; pipeline processing; query processing

Identifiers: distributed **inverted index**; **Web pages**; pipelining technique; core index-building system; index construction time; storage scheme; **inverted files**; embedded database system; global statistics;

distributed full-text index

Class Codes: C7250L (Non-bibliographic retrieval systems); C7210N (Information networks); C4240P (Parallel programming and algorithm theory); C6120 (File organisation)

Copyright 2002, IEE

7/5/6 (Item 4 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2004 Institution of Electrical Engineers. All rts. reserv.

7129085 INSPEC Abstract Number: C2002-01-6180-040

Title: The active guidebook

Author(s): Boyd, T.; Robinson, P.

Author Affiliation: Comput. Lab., Cambridge Univ., UK

Conference Title: New Developments in Distributed Applications and Interoperable Systems. IFIP TC6/WG6.1 Third International Working Conference on Distributed Applications and Interoperable Systems p.55-66

Editor(s): Zielinski, K.; Geihs, K.; Laurentowski, A.

Publisher: Kluwer Academic Publishers, Norwell, MA, USA

Publication Date: 2001 Country of Publication: USA xii+322 pp.

ISBN: 0 7923 7481 9 Material Identity Number: XX-2001-02120

Conference Title: Proceedings of International Working Conference on Distributed Applications and Interoperable Systems

Conference Sponsor: IFIP ,

Conference Date: 17-19 Sept. 2001 Conference Location: Krakow, Poland

Language: English Document Type: Conference Paper (PA)

Treatment: Applications (A); General, Review (G)

Abstract: The active guidebook is a context-aware information management system that uses a combination of spatial and keyword indexing to retrieve data. The system has three principal components: a new document description language extends HTML to include facilities for tagging with spatial locations. Retrieval uses two separate indexes - a segment tree is used for spatial indexing and an **inverted file** is used for keyword indexing. A user interface allows queries involving keywords and location data to be expressed, and presents their results. The system has been evaluated with the implementation of an interactive guidebook. The test data was drawn from existing **Web pages** describing the city of Cambridge in England, which were augmented with spatial information. A GPS system is used to provide the default location information for retrieval, but can be overridden with explicit coordinates. (10 Refs)

Subfile: C

Descriptors: Global Positioning System; hypermedia markup languages; indexing; information retrieval systems; mobile computing; tree searching; user interfaces

Identifiers: active guidebook; context-aware information management system; spatial indexing; keyword indexing; data retrieval; document description language; HTML; tagging; segment tree; **inverted file** ; user interface; interactive guidebook; **Web pages** ; Cambridge; GPS; default location information

Class Codes: C6180 (User interfaces); C7250 (Information storage and retrieval); C6120 (File organisation); C6150N (Distributed systems software); C6130M (Multimedia)

Copyright 2001, IEE

7/5/7 (Item 5 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2004 Institution of Electrical Engineers. All rts. reserv.

6498819 INSPEC Abstract Number: C2000-03-7250N-011

Title: Needles and Haystacks: a search engine for personal information collections

Author(s): de Kretser, O.; Moffat, A.

Author Affiliation: Dept. of Comput. Sci. & Software Eng., Melbourne Univ., Vic., Australia

Conference Title: Proceedings 23rd Australasian Computer Science Conference. ACSC 2000 (Cat. No.PR00518) p.58-65

Editor(s): Edwards, J.
Publisher: IEEE Comput. Soc, Los Alamitos, CA, USA
Publication Date: 1999 Country of Publication: USA ix+249 pp.
ISBN: 0 7695 0518 X Material Identity Number: XX-2000-00244
U.S. Copyright Clearance Center Code: 0 7695 0518 X/99/\$10.00
Conference Title: Proceedings 23rd Australasian Computer Science Conference. ACSC 2000
Conference Sponsor: Australian Nat Univ.; Univ. Canberra; Univ. Coll., Australian Defence Force Acad.; Linuxcare; ACSys
Conference Date: 31 Jan.-3 Feb. 2000 Conference Location: Canberra, ACT, Australia
Language: English Document Type: Conference Paper (PA)
Treatment: Applications (A); Practical (P)
Abstract: Information retrieval systems can be partitioned into two main classes: large-scale systems that make use of an **inverted index** or some other auxiliary data structure, intended for massive volumes of data; and the small-scale systems based upon sequential pattern matching that most computer users employ when hunting for missing **email** and news items. In this paper we describe a hybrid approach that offers the ranked queries and similarity matching of a genuine information retrieval system, but does so without any need for an index to be precomputed. This software tool, which we call **seft**, offers performance that in a retrieval effectiveness sense matches conventional information retrieval systems, and in a resource efficiency sense, while considerably slower than **grep**-like tools, is fast enough to be useful on hundreds of megabytes of text. (19 Refs)
Subfile: C
Descriptors: pattern matching; search engines
Identifiers: search engine; personal information collections; information retrieval systems; large-scale systems; **inverted index**; auxiliary data structure; small-scale systems; sequential pattern matching; **email**; news items; hybrid approach; ranked queries; similarity matching; software tool; **seft**; **grep**-like tools; Needles and Haystacks
Class Codes: C7250N (Search engines); C7210N (Information networks)
Copyright 2000, IEE

7/5/8 (Item 6 from file: 2)
DIALOG(R)File 2:INSPEC
(c) 2004 Institution of Electrical Engineers. All rts. reserv.

6442069 INSPEC Abstract Number: C2000-01-7240-022
Title: Dynamic email organization via relevance categories
Author(s): Mock, K.
Author Affiliation: Intel Corp., USA
Conference Title: Proceedings 11th International Conference on Tools with Artificial Intelligence p.399-405
Publisher: IEEE Comput. Soc, Los Alamitos, CA, USA
Publication Date: 1999 Country of Publication: USA xvii+446 pp.
ISBN: 0 7695 0456 6 Material Identity Number: XX-1999-03348
U.S. Copyright Clearance Center Code: 0 7695 0456 6/99/\$10.00
Conference Title: Proceedings 11th International Conference on Tools with Artificial Intelligence. TAI 99
Conference Sponsor: IEEE Comput. Soc.; AAAI Soc.; Virtual Intelligence Task Force; BU-CIS Center
Conference Date: 9-11 Nov. 1999 Conference Location: Chicago, IL, USA
Language: English Document Type: Conference Paper (PA)
Treatment: Experimental (X)
Abstract: Many researchers have proposed classification systems that automatically classify **email** in order to reduce information overload. However, none of these systems are in use today. This paper examines some of the problems with classification technologies and proposes Relevance Categories as a method to avoid some of these problems. In particular, the dynamic nature of **email** categories, the cognitive overhead, required training categories, and the high costs of classification errors are hurdles for many classification algorithms. Relevance Categories avoid some of these problems through their simplicity; they are merely relevance-ranked lists of **email** messages that are similar to a set of query messages. by displaying messages as the result of a dynamic query in

lieu of fixed categories, we hypothesize that users will be less sensitive to errors using the Relevance Categories scheme than to errors using a fixed categorization scheme. To study the effectiveness of the Relevance Categories concept, we devised a performance metric for relevance ranking and used it to test an **inverted index** implementation on the Reuter-21578 test collection. The promising test results indicate the need for further work. (9 Refs)

Subfile: C

Descriptors: classification; **electronic mail** ; relevance feedback

Identifiers: **email** organization; relevance categories; classification systems; cognitive overhead; training categories; performance metric; relevance ranking; **inverted index**

Class Codes: C7240 (Information analysis and indexing); C7250R (Information retrieval techniques)

Copyright 1999, IEE

7/5/9 (Item 7 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2004 Institution of Electrical Engineers. All rts. reserv.

5610359 INSPEC Abstract Number: C9708-6160Z-001

Title: **Data warehouse processing and query by mail**

Author(s): Gordon, K.

Journal: APL Quote Quad vol.27, no.1 p.39-44

Publisher: ACM,

Publication Date: Sept. 1996 Country of Publication: USA

CODEN: APLQD9 ISSN: 0163-6006

SICI: 0163-6006(199609)27:1L:39:DWPQ;1-1

Material Identity Number: B763-97002

Language: English Document Type: Journal Paper (JP)

Treatment: Practical (P)

Abstract: Much effort in IS is going into creating data warehouses. These are stores of data periodically extracted from older legacy applications, converted to common standards and made accessible for user analysis. The warehouse acts as a WORM (Write Once, Read Many times) storage. Where the extract and transfer is performed nightly, they provide access to what is termed "near operational" data and can be used to replace much of the existing reporting. In other cases they are used to store mostly historical data for analysis of trends, market impact, financial status and so on. While often implemented with a variety of different clean up tools, languages, database products and query tools, this article describes an implementation done almost entirely with APL. It includes a query capability termed "query by mail" which enables anyone with access to **e-mail** to send queries to the warehouse and receive responses or extracts of data by return mail. The "query" includes customized analysis of field content to allow identification of fields and records containing invalid data. Built upon a proprietary **inverted file** system, it provides rapid response to user queries and little load on the server system. (0 Refs)

Subfile: C

Descriptors: APL; business data processing; client-server systems; data integrity; **electronic mail** ; query processing; very large databases

Identifiers: data warehouse processing; query by mail; information systems; legacy applications; client server system; user analysis; WORM; Write Once Read Many; near operational data; historical data; market impact ; financial status; clean up tools; database; query tools; APL; **e-mail** ; customized analysis; invalid data; **inverted file** system

Class Codes: C6160Z (Other DBMS); C7100 (Business and administration); C6140D (High level languages)

Copyright 1997, IEE

7/5/10 (Item 8 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2004 Institution of Electrical Engineers. All rts. reserv.

3513263 INSPEC Abstract Number: C90003609

Title: **Electronic filing and retrieval: developments in full text**

retrieval systems

Author(s): Pritchard, J.

Author Affiliation: Nat. Comput. Centre Ltd., Manchester, UK

Conference Title: OIS International 1989. Proceedings of the Sixth Annual Conference on Optical Information Systems p.207-15

Publisher: Meckler, London, UK

Publication Date: 1989 Country of Publication: UK 248 pp.

Conference Sponsor: Meckler; Cimtech

Conference Date: 15-17 May 1989 Conference Location: London, UK

Language: English Document Type: Conference Paper (PA)

Treatment: General, Review (G)

Abstract: Full text retrieval (FTR) technology has evolved very significantly during the late 1980s. Organisations are incorporating an FTR system into their corporate IT strategy, and integrating it with word processing, **electronic mail**, micro-mainframe links, integrated office systems and document image processing systems. The paper discusses the future requirements of FTR systems covering aspects such as recall and precision, **inverted file** indexing, optical character recognition (OCR) technology, multi-lingual capabilities, and automatic hypertext generator systems. (8 Refs)

Subfile: C

Descriptors: computerised picture processing; hypermedia; information retrieval systems

Identifiers: full text retrieval technology; electronic filing; electronic retrieval; EFR; FTR system; corporate IT strategy; word processing; **electronic mail**; micro-mainframe links; integrated office systems; document image processing systems; **inverted file** indexing; optical character recognition; multi-lingual capabilities; automatic hypertext generator systems

Class Codes: C7250L (Non-bibliographic systems); C7100 (Business and administration)

7/5/11 (Item 1 from file: 144)

DIALOG(R)File 144:Pascal

(c) 2004 INIST/CNRS. All rts. reserv.

16899475 PASCAL No.: 04-0561910

Index construction for linear categorisation

Twelfth ACM international conference on information & knowledge management : New Orleans LA, 3-8 November 2003

SHANKS Vaughan R; WILLIAMS Hugh E

FRIEDER Ophir, ed; HAMMER Joachim, ed; QHERSHI Sajda, ed; SELIGMAN Len, ed

School of Computer Science and Information Technology, RMIT University, GPO Box 2476V, Melbourne 3001, Australia

ACM. SIGIR special interest group on information retrieval, United States ; ACM. SIGMIS special interest group on management information systems, United States

CIKM 2003 : conference on information and knowledge management, 12 (New Orleans LA USA) 2003-11-03

2003 334-341

Publisher: Association for Computing Machinery, New York NY

ISBN: 1-58113-723-0 Availability: INIST-Y 38186; 354000117910020460

No. of Refs.: 35 ref.

Document Type: C (Conference Proceedings) ; A (Analytic)

Country of Publication: United States

Language: English

Categorisation is a useful method for organising documents into subcollections that can be browsed or searched to more accurately and quickly meet information needs. On the **Web**, category-based **portals** such as Yahoo! and DMOZ are extremely popular: DMOZ is maintained by over 56,000 volunteers, is used as the basis of the popular Google directory, and is perhaps used by millions of users each day. Support Vector Machines (SVM) is a machine-learning algorithm which has been shown to be highly effective for automatic text categorisation. However, a problem with iterative training techniques such as SVM is that during their learning or training phase, they require the entire training collection to be held in

main-memory; this is infeasible for large training collections such as DMOZ or large news wire feeds. In this paper, we show how inverted indexes can be used for scalable training in categorisation, and propose novel heuristics for a fast, accurate, and memory efficient approach. Our results show that an index can be constructed on a desktop workstation with little effect on categorisation accuracy compared to a memory-based approach. We conclude that our techniques permit automatic categorisation using very large training collections, vocabularies, and numbers of categories.

English Descriptors: **Index** (documentation); **Inverted file** ;
Construction; Performance evaluation

Broad Descriptors: Information retrieval; Automatic classification;
Recherche information; Classification automatique; Busqueda informacion;
Clasificacion automatica

French Descriptors: Index; Fichier inverse; Construction; Evaluation
performance; SVM (Support Vector Machine)

Classification Codes: 001A01F03; 205

Copyright (c) 2004 INIST-CNRS. All rights reserved.

7/5/12 (Item 2 from file: 144)
DIALOG(R) File 144:Pascal
(c) 2004 INIST/CNRS. All rts. reserv.

16343199 PASCAL No.: 03-0508950

Optimization of restricted searches in Web directories using hybrid data structures

Advances in information retrieval : Pisa, 14-16 April 2003

CACHEDA Fidel; CARNEIRO Victor; GUERRERO Carmen; VINA Angel

SEBASTIANI Fabrizio, ed

Department of Information and Communications Technologies, Facultad de Informatica, Campus de Elvifia s/n, 15.071, Coruna, Spain

ECIR 2003 : European conference on IR research, 25 (Pisa ITA) 2003-04-14

Journal: Lecture notes in computer science, 2003, 2633 436-451

ISBN: 3-540-01274-5 ISSN: 0302-9743 Availability: INIST-16343;

354000108528470310

No. of Refs.: 18 ref.

Document Type: P (Serial); C (Conference Proceedings) ; A (Analytic)

Country of Publication: Germany

Language: English

The need of efficient tools in order to manage, retrieve and filter the information in the WWW is clear. **Web directories** are taxonomies for the classification of Web documents. These kind of information retrieval systems present a specific type of search where the document collection is restricted to one area of the category graph. This paper introduces a specific data architecture for **Web directories** that improves the performance of restricted searches. That architecture is based on a hybrid data structure composed of an **inverted file** with multiple embedded signature files. Two variants are presented: hybrid architecture with total information and with partial information. This architecture has been analyzed by means of developing both variants to be compared with a basic model. The performance of the restricted queries was clearly improved, especially the hybrid model with partial information, which yielded a positive response under any load of the search system.

English Descriptors: Information retrieval; World wide **web** ; **Directory** ;
Classification; Data structure; Hybrid model; Search system; System
description; Hybrid system; **Inverted file** ; Filter; Performance
evaluation; System architecture; Signing; **Internet directory** ; Partial
information; Total information

French Descriptors: Recherche information; Reseau web; Repertoire;
Classification; Structure donnee; Modele hybride; Systeme recherche;
Description systeme; Systeme hybride; Fichier inverse; Filtre; Evaluation

performance; Architecture systeme; Signature; Annuaire internet;
Information partielle; Information totale

Classification Codes: 001A01F03; 001A01E02C1; 205

Copyright (c) 2003 INIST-CNRS. All rights reserved.

7/5/14 (Item 4 from file: 144)
DIALOG(R) File 144:Pascal
(c) 2004 INIST/CNRS. All rts. reserv.

12672302 PASCAL No.: 96-0372386

World wide web-based information storage and retrieval

O'KANE K C

Department of Computer Science, University of Northern Iowa, Cedar Falls,
IA 50613-0507, United States

Journal: Online & CDROM review, 1996, 20 (1) 11-20

ISSN: 1353-2642 Availability: INIST-17093; 354000044400620020

No. of Refs.: 10 ref.

Document Type: P (Serial) ; A (Analytic)

Country of Publication: United Kingdom

Language: English

This paper describes the design and implementation of a system for computer generation of linked HTML documents to support information retrieval and hypertext applications on the World Wide Web. The approach is based on work by Salton and others, but extends the concept to be compatible with the World Wide Web browser environment by adding an interactive indexing technique that is well suited to the mouse-based point-and-shoot input common to windowed browsers. The system does not require text query input, nor any client or host processing other than hypertext linkage. The goal of this work is to construct a fully automatic system in which original text documents are read and processed by a computer program that generates HTML files, which can be used immediately by Web browsers to search and retrieve the original documents. Thus, a user with a large collection of information—for instance, newspaper articles—can feed these documents to the program described here and produce directly, without further human intervention, the necessary files to establish World Wide Web **home** and related **pages**, to support interactive retrieval and distribution of the original documents.

English Descriptors: Bibliographic search; Query; Natural language;

Hypertext; Document access; Document processing; Automated processing;

Vector method; Automatic indexing; **Inverted file**; Construction;

Computer program; Information retrieval software; System design;

Implementation; World Wide Web; HTML document

Broad Descriptors: Documentation data processing; Informatique documentaire
; Informacion documental

French Descriptors: Recherche bibliographique; Question documentaire;

Langage naturel; Hypertexte; Acces document; Traitement document;

Traitement automatise; Methode vectorielle; Indexation automatique;

Fichier inverse; Construction; Programme ordinateur; Logiciel

documentaire; Conception systeme; Implementation; Salton, G.; HTML

(HyperText Markup Language); World Wide Web; Document HTML

Classification Codes: 001A01F03; 205

?

?ds

Set	Items	Description
S1	2499635	ADVERTIS? OR ADVERTIZ? OR AD OR ADS OR PROMOTION? OR ADVERT? T? ? OR COMMERCIAL()MESSAGE?
S2	409	(INVERT?) (2N) (INDEX? ? OR INDICES OR FILE OR FILES OR FI- LING? OR LIST OR LISTS OR LISTING? OR STRUCTURE?)
S3	629726	(INTERNET OR WEB OR ONLINE OR ON()LINE OR HOME) (2N) (PAGE OR PAGES OR SITE OR SITES OR PORTAL? OR DIRECTOR?)
S4	961879	(E OR ELECTRONIC OR DIGITAL OR VIRTUAL) (1W) (MAIL??? OR M- ESSAG??? OR CORRESPOND?) OR EMAIL???? OR (INTERNET OR ON()LI- NE OR ONLINE OR WEB) (1W) MAIL???? OR MIME OR SMTP OR POP(1N)- MAIL
S5	8	S1 (S)S2
S6	7	RD (unique items)
S7	24	S1 AND S2 AND (S3 OR S4)
S8	22	RD (unique items)
S9	18	S8 NOT S6
S10	10	S2 (S) (S3 OR S4)
S11	8	RD (unique items)
S12	3	S11 NOT (S6 OR S9)

?show files

File 635:Business Dateline(R) 1985-2005/Jan 06
(c) 2005 ProQuest Info&Learning

File 570:Gale Group MARS(R) 1984-2005/Jan 06
(c) 2005 The Gale Group

File 47:Gale Group Magazine DB(TM) 1959-2005/Jan 06
(c) 2005 The Gale group

File 387:The Denver Post 1994-2005/Jan 05
(c) 2005 Denver Post

File 471:New York Times Fulltext 1980-2005/JAN 06
(c) 2005 The New York Times

File 492:Arizona Repub/Phoenix Gaz 19862002/Jan 06
(c) 2002 Phoenix Newspapers

File 494:St LouisPost-Dispatch 1988-2005/Jan 03
(c) 2005 St Louis Post-Dispatch

File 498:Detroit Free Press 1987-2004/Dec 30
(c) 2005 Detroit Free Press Inc.

File 631:Boston Globe 1980-2004/Dec 31
(c) 2005 Boston Globe

File 633:Phil.Inquirer 1983-2004/Dec 31
(c) 2005 Philadelphia Newspapers Inc

File 638:Newsday/New York Newsday 1987-2005/Jan 04
(c) 2005 Newsday Inc.

File 640:San Francisco Chronicle 1988-2005/Jan 06
(c) 2005 Chronicle Publ. Co.

File 641:Rocky Mountain News Jun 1989-2005/Jan 06
(c) 2005 Scripps Howard News

File 702:Miami Herald 1983-2005/Jan 05
(c) 2005 The Miami Herald Publishing Co.

File 703:USA Today 1989-2005/Jan 05
(c) 2005 USA Today

File 704:(Portland)The Oregonian 1989-2004/Dec 31
(c) 2005 The Oregonian

File 713:Atlanta J/Const. 1989-2005/Jan 06
(c) 2005 Atlanta Newspapers

File 714:(Baltimore) The Sun 1990-2004/Dec 31
(c) 2005 Baltimore Sun

File 715:Christian Sci.Mon. 1989-2005/Jan 06
(c) 2005 Christian Science Monitor

File 725:(Cleveland)Plain Dealer Aug 1991-2005/Jan 05
(c) 2005 The Plain Dealer

File 735:St. Petersburg Times 1989- 2005/Jan 05
(c) 2005 St. Petersburg Times

File 476:Financial Times Fulltext 1982-2005/Jan 06
(c) 2005 Financial Times Ltd

File 477:Irish Times 1999-2005/Jan 06
(c) 2005 Irish Times

File 710:Times/Sun.Times(London) Jun 1988-2005/Jan 05

(c) 2005 Times Newspapers

File 711:Independent(London) Sep 1988-2005/Jan 05

(c) 2005 Newspaper Publ. PLC

File 756:Daily/Sunday Telegraph 2000-2005/Jan 06

(c) 2005 Telegraph Group

File 757:Mirror Publications/Independent Newspapers 2000-2005/Jan 05

(c) 2005

?

?ts9/3,k/14

9/3,K/14 (Item 10 from file: 47)
DIALOG(R)File 47:Gale Group Magazine DB(TM)
(c) 2005 The Gale group. All rts. reserv.

04713411 SUPPLIER NUMBER: 19203323 (USE FORMAT 7 OR 9 FOR FULL TEXT)
Day tripping to Internet World. (Fall 1996 trade show)
Bjorner, Susanne
Searcher, v5, n2, p50(12)
Feb, 1997
ISSN: 1070-4795 LANGUAGE: English RECORD TYPE: Fulltext
WORD COUNT: 9526 LINE COUNT: 00813

... by providing, a service (including but not limited to finding information). Services provided might include **advertising** opportunities, a research service (e.g., Find/SVP, or travel services. * Technology: Products that are...

...searchers. * Intranet Development: Products that assist in the development or maintenance of an Intranet or **Internet Web site**, such as InmagicB/Text Webserver.

These are loose categorizations only; several products cover more than...

...you need.

OK, here goes....

A Business Compass

Does the world really need a new **online Web directory**? Probably not, but it certainly could use a better one or two. Quite possibly, A...

...Business Compass LLB executive vice president, he took the service live in mid-December. This **Internet directory** and analysis service pre-screens and profiles leading Web-based sources of business information and...www.adobe.com/).

Adobe Pagemill 2.0 for the Macintosh, which enhances and speeds up **Web page** design, and Adobe's Photoshop 4.0 for Windows and Macintosh platforms, which supports digital...

...may select an interface language from 18 different languages, and the browser will accurately display **Web pages** authored in any of over 90 world languages. Arabic and Hebrew characters display properly in...

...and multimedia content.

The Tango 2.5 browser is available for downloading from the Alis **Web site** (<http://www.alis.com>). They let you "test-drive" it free for up to 30...

...0, which will integrate Tango Mail, a feature allowing easy, automatic creation and viewing of **e-mail** in dozens of languages.

These and other communication and translation services and tools are part...

...surcharge.

Citizen 1 Software, Inc.

Continuing the trend toward making the retrieval of information from **Web sites** at least as easy and fast as locating information from print-based resources, Citizen 1...

...annual cost-per-seat basis.

Citysurf

Virtual Media Services, of Tulsa, Oklahoma, demonstrated and sought **advertisers** for a nationwide business Net directory. Citysurf users can find a business in over 90...the national Bigbook, Big Yellow, US WEST, and Athand directories. They have especially targeted local **advertising** and local sites. Citysurf offers **advertisers** the opportunity to buy banner **ads** or to be listed in the "Top 10" of a specific city in their business ...

...workflow, software to colleges and universities. hyperlinks to more than 2,000 college and university **home pages** .

Direcpc

DirecPC, from Hughes Network Services, is the U.S. domestic alternative for satellite Internet...

...stores began retailing DirecPC in California in October, and if you go to DirectPC's **Web site** (<http://www.direcpc.com>), you can get a listing of retail suppliers in your area...

...I'll continue to hope.

Dun & Bradstreet

DAB and Lycos, Inc. announced the co-branded **Web site** CompaniesOnline (<http://www.companiesonline.com>), a free business-to-business directory featuring detailed information on...

...search criteria; 2) a detailed company information page; and 3) a page with company-supplied **advertising** information. The site includes banner **advertising** .

D&B brings company-specific information from its database of more than 40 million companies...

...which contains detailed company data supplied by DAB, users can link to the company's **Web site** or a page containing **advertising** information supplied by the company. **Advertisers** pay an annual subscription fee for inclusion. Business users can also obtain company information by linking directly to D&B's **Web site** and purchasing a \$20 Business Background Report on the company by using a credit card...

...the business name and location (city and state); the company information page includes mailing address, **e - mail** address, D&B D-U-N-S Number, telephone number, trade-style name, company size...

...legal status, parent company name, contact name and title, and company URL. Information on the **advertising** page varies by company and requires registration with Lycos.

D&B also demonstrated its Marketing...

...Excite, Hotbot, Infoseek, Infoseek Ultra, Lycos, Magellan, Open Text Index, Webcrawler, and Yahoo!. Reports are **e - mailed** to requesters and can be kept on file as an indication of the status of...Family & Relationships, and Health. Unlike typical Web search engines, FINDOUT's libraries direct users to **sites** on commercial **online** services and offline articles, books, organizations, guides, videos, software, and CD-ROMS.

Find/SVP is...

...discussion with prospective sponsors and with leading membership and subscription-based service organizations, ISPS, and **Web sites** interested in providing on-demand answers as a value-added customer benefit. Find/SVP is...

...FINDOUT button" on their own sites as a service promoting brand loyalty and repeat visits. **Advertising** and sponsorship opportunities and 800- or 900-number telephone access options are also in the...

...Forefront is in the forefront. Their Webwhacker offline browsing product created the standard for downloading **Web sites** to a user's hard drive. Webseeker, now in an enhanced version 2.2, runs...

...offline. Keyword highlighting enables users to quickly locate searched words within the text of the **Web pages** opened. The enhanced version includes a scheduling interface and allows the setting of download control limits.

At COMDEX, Forefront announced Webprinter 2.0, an application that allows users to turn **Web pages** into attractive, double-sided, hardcopy booklets. Webprinter intercepts **Web pages** on the way to the printer, automatically reducing, rotating, and realigning them to print as...

...CompuServe, America Online, Prodigy, Mosaic, and other browsers. Trial versions are available at the Forefront **Web site** [Chttp://www.ffg.com](http://www.ffg.com)). Webprinter is based on technology found in Forefront's more powerful...

...navigate, and organize into a seamless interface.

GRIT

If you have tired of reading about **Internet sites**, you can hear about them instead on GRIT (Gould Resources & Internet Telecommunications), "the world's...

...day, seven days a week, via RealAudio and StreamWorks, at <http://www.grit.com>. The **Web site** features live talk about sports, exercise, music, technology, politics, etc., and also reviews **Web sites**. It also contains a search engine, a digital photograph gallery, and a link to GRIT's CU-SeeMe reflector site. Listeners may visit other **Web sites** while the shows play in the background.

HDS Network Systems

At Internet World, I had...the Fly Conference. The software makes it easy for users to post messages on a **Web site** in an organized manner, in either public or private modes. Conference supports private conferences with...

...intelligently crawl more than 10 million Web documents per day and is said to index **Web site** content up to three times faster than any other technology available. HotBot received the PC...

...the formation of the NewsPage Network (NPN), which makes daily customized news available to other **Web sites**. Designed as a traffic builder for specialized content sites and a wider distribution area for...

...future than they do now. You can already see traces of NewsPage on the following **Web sites**: MSNBC, Quicken Financial Network, InfoSeek, Achoo Online Healthcare, Kleiner Perkins Caulfield & Byers, All Things...

...bed and breakfast in Northern California. the new Ultrasmart not only provides a listing of **Web pages** in response to a query, but also finds material on related topics, such as Hotels...

...World. One of those was InMagic, Inc., who announced shipment of DB/Text WebServer, the **Internet** and Intranet **site** management version of the venerable and versatile InMagic database/textbase software. A Word Wheel feature...

...words in a DB/Text WebServer database.

Traditional online searchers may recognize this as an **inverted file**, ..ROOTing, or Expanding, but the word wheel terminology puts a nice spin on the old...Internet or Intranet." Possible implementations of Messenger include clickable news headlines. headlines hotlinked to corresponding **Web pages**, **advertising** and "billboards" that register clicks to indicate interest for initiating sales processes; and priority messages" that deliver highly urgent communications and links to **Web pages** for more information or action.

The Internet Company also produces NewsSpace, a search and crawl tool for **Web sites** and provides consulting for Internet-related design, development, programming, and marketing. Its client list includes...

...Hewlett-Packard Company and Live Picture, Inc. jointly announced the arrival of an Imaging for **Internet Web site**, @ a public beta site demonstrating technology for viewing, sharing, and printing high-resolution images from the **Internet**. The **site** (<http://www.image.hp.com>) allows the display of photo-rich content and enables users to download the imaging for Internet technology. Currently, the **Web page** includes content samples from Corbis Corporation Photo Collection (selections from award winning photographers, museums, and...

...1997 issue, NC World grows to full size as a stand-alone electronic publication, an **advertiser**-supported monthly magazine with mid-month updates. Access is free. Subscribers who fill out a detailed demographic form will receive **e-mail** alerts of new articles. (<http://www>).

ncworldmag.com).

NetCarta Corporation

Netcarta introduced version 2.0 of its WebMapper content management software, which creates an object-based, high-level map of **Web sites** for WebMasters. NetCarta WebMapper 2.0 automates most common site management tasks required in the maintenance of **Internet** and Intranet **sites**.

WebMapper creates a Tree View, which provides an orderly and easy-to-read hierarchical outline...

...new "Cyberbolic" view. This Web-like gestalt view shows site objects fanning out from the **home page**, then the next level of pages, and so on, effectively giving a bird's eye...

...average amount of time it takes end users to access a selection of high-traffic **Web sites**.

ZDNet.Sweep should debut in the first quarter of 1997, published in three places: the ZD Net **Web site** (<http://www.zdnet.com>), where it will update several times each day, Inter@ctive Week...

...compare their daily access and retrieval times to the average performance of selected high-traffic **Web sites**.

The net.Sweep polling system is based on a distributed network across the United States...

...are established in major cities and through multiple Internet Service Providers (ISPs). The system accesses **Web sites** in the ZDNet.Sweep index over 500 times per day, seven days a week, resulting...introduced Netscape Communicator, a comprehensive set of Internet and Intranet component software that integrates open **e-mail**, groupware, editing, calendaring, and browsing tools to allow users to easily communicate, share, and access...

...Navigator (the popular Web browser), Netscape Messenger @allowing the composition, sending and receiving of encrypted **e-mail**, using open-standard based mail), Netscape Collabra (facilitating collaboration with co-workers and leveraging corporate...

...team; posted notes can have an expiration date and urgency markings to denote importance); and **E-mail** Notification (users can subscribe to get notices via **e-mail** when something significant, such as an object, message, or task, changes in the project.).

PaperClip...

...Microsoft Internet Explorer 3.0, and Spyglass Mosaic 2.10 or above. WebClip automatically checks **Web pages**. identifies whether any changes have occurred on a page, and sends the new content to a user's PC. Users can then display the same **Web pages** at hard drive speed, with all graphics and hotlinks intact. Users can schedule unattended monitoring...

...to maintain collections of Web content even when the content no longer resides on the **Web site** itself.

PaperClip has expanded its retail distribution to individuals and enterprises via numerous chains and...

...spiders. AT1 also allows an users to create intelligent agents to the search newsgroups and **Web sites**, with **e-mail** notification of results.

In a statement that will ring true to experienced searchers, PLS founder...

...Virtual Address Book if you lose yours. Best of all. you can access PlanetAll by **Web**. **e-mail**. touch-tone phone, or fax. Also offered is a Special Occasions Reminder. which alerts members...Spectrum and Find/SVP), QUALCOMM's Eudora can rightfully claim to be the leader in **Internet e-mail** software. At Internet World, QUALCOMM announced the availability of Eudora Light 3.0.1 for...

...download at <http://www.eudora.com>.

Eudora Light is known as an easy-to-use e - mail solution for new Internet users. Many ISPs distribute it for free. Major new features of Eudora Light 3.0.1. include basic filters to organize e - mail automatically, an enhanced Find dialog, mail server interactive control, drag and drop capabilities for file...

...features (not in Eudora Light 3.0.1 freeware) include a user controllable toolbar, multiple e - mail accounts., enhanced filters, including auto-reply and forwarding, spell checking., composition of stylized text, expanded...

...features to its travel-related services. Launched a year ago, Travelocity has become a popular Web site for do-it-yourself travel arrangers. Users can choose from among 50 car rental companies...the speed and cost-effectiveness of electronic delivery. The Posta user enters a recipient's e - mail address (or a mailing list) and selects a document to send. Tumbleweed Posta places the document on a powerful server, which sends an e - mail notification to the recipient. When recipients receive notification of a document's arrival through their e - mail client software, they simply click on an address to receive the document.

Tumbleweed Posta users...

...deliver an appropriate viewer or plug-in with the document, when needed. Anyone with an Internet e - mail account can receive Posta documents, regardless of their recipient e - mail system or hardware and software capabilities. Tumbleweed Posta users can track documents through every step ...

...a photographic memory of Web sessions. While you surf, ZooWorks records and indexes all the Web pages and documents you visit. This information, along with the content of the HTML document, is...for searchers. Intranet: Products that assist in the development or maintenance of an Intranet or Internet Web site .

How to Talk to Internet World Exhibitors:

* Tell them "I'm happy to pay for..."

?

?t: sl2/3,k/1

12/3,K/1 (Item 1 from file: 570)
DIALOG(R)File 570:Gale Group MARS(R)
(c) 2005 The Gale Group. All rts. reserv.

02081369 Supplier Number: 74649324 (USE FORMAT 7 FOR FULLTEXT)
How a Search Engine Works. (Internet/Web/Online Service Information)
Liddy, Elizabeth
Searcher, v9, n5, p38
May, 2001
ISSN: 1070-4795
Language: English Record Type: Fulltext
Document Type: Magazine/Journal; Professional
Word Count: 4015

... simply standardize the multiple formats encountered when deriving documents from various providers or handling various **Web sites**. The steps serve to merge all the data into a single consistent data structure that...

...later steps of document processing. Step two is important because the pointers stored in the **inverted file** will enable a system to retrieve various sized units -- either site, page, document, section, paragraph...
?

?qs

Set	Items	Description
S1	6843167	ADVERTIS? OR ADVERTIZ? OR AD OR ADS OR PROMOTION? OR ADVERT- T? ? OR COMMERCIAL()MESSAGE?
S2	1800	(INVERT?) (2N) (INDEX? ? OR INDICES OR FILE OR FILES OR FI- LING? OR LIST OR LISTS OR LISTING? OR STRUCTURE?)
S3	7309091	(INTERNET OR WEB OR ONLINE OR ON()LINE OR HOME) (2N) (PAGE OR PAGES OR SITE OR SITES OR PORTAL? OR DIRECTOR?)
S4	4079577	(E OR ELECTRONIC OR DIGITAL OR VIRTUAL) (1W) (MAIL??? OR M- ESSAG??? OR CORRESPOND?) OR EMAIL???? OR (INTERNET OR ON()LI- NE OR ONLINE OR WEB) (1W) MAIL???? OR MIME OR SMTP OR POP(1N)- MAIL
S5	29	S1(S)S2
S6	24	RD (unique items)
S7	17	S6 NOT PY>2003
S8	105	S1 AND S2 AND S3:S4
S9	14	S8 AND S2(S) (S3 OR S4)
S10	8	RD (unique items)
S11	8	S10 NOT S7
S12	32	S2(S) (S3:S4)
S13	22	RD (unique items)
S14	14	S13 NOT (S11 OR S7)
S15	14	S14 NOT PY>2003

?show files

File 15:ABI/Inform(R) 1971-2005/Jan 05
(c) 2005 ProQuest Info&Learning

File 9:Business & Industry(R) Jul/1994-2005/Jan 05
(c) 2005 The Gale Group

File 610:Business Wire 1999-2005/Jan 06
(c) 2005 Business Wire.

File 810:Business Wire 1986-1999/Feb 28
(c) 1999 Business Wire

File 275:Gale Group Computer DB(TM) 1983-2005/Jan 06
(c) 2005 The Gale Group

File 476:Financial Times Fulltext 1982-2005/Jan 06
(c) 2005 Financial Times Ltd

File 624:McGraw-Hill Publications 1985-2005/Jan 05
(c) 2005 McGraw-Hill Co. Inc

File 621:Gale Group New Prod.Annou.(R) 1985-2005/Jan 06
(c) 2005 The Gale Group

File 636:Gale Group Newsletter DB(TM) 1987-2005/Jan 06
(c) 2005 The Gale Group

File 613:PR Newswire 1999-2005/Jan 05
(c) 2005 PR Newswire Association Inc

File 813:PR Newswire 1987-1999/Apr 30
(c) 1999 PR Newswire Association Inc

File 16:Gale Group PROMT(R) 1990-2005/Jan 06
(c) 2005 The Gale Group

File 160:Gale Group PROMT(R) 1972-1989
(c) 1999 The Gale Group

File 634:San Jose Mercury Jun 1985-2004/Dec 31
(c) 2005 San Jose Mercury News

File 148:Gale Group Trade & Industry DB 1976-2005/Jan 06
(c) 2005 The Gale Group

File 20:Dialog Global Reporter 1997-2005/Jan 06
(c) 2005 The Dialog Corp.

File 696:DIALOG Telecom. Newsletters 1995-2005/Jan 05
(c) 2005 The Dialog Corp.

File 647:CMP Computer Fulltext 1988-2005/Dec W3
(c) 2005 CMP Media, LLC

File 674:Computer News Fulltext 1989-2004/Dec W2
(c) 2004 IDG Communications

?

?t s11/3,k/2,3,6

11/3,K/2 (Item 2 from file: 15)
DIALOG(R)File 15:ABI/Inform(R)
(c) 2005 ProQuest Info&Learning. All rts. reserv.

02224682 80461446

Let's stop the trash trucks

Quint, Barbara

Information Today v18n8 PP: 8-14 Sep 2001

ISSN: 8755-6286 JRNL CODE: IFT

WORD COUNT: 2655

...ABSTRACT: be considered that would protect everyone. Online commercial services and searchable database archives on publisher **Web sites** should continue to maintain the **inverted file index** terms and tags that identify material barred from full-text delivery by the Tasini decision. The **inverted file** indexes belong to the host services, and they could continue to identify relevant articles, at...

...TEXT: access to all New York Times book reviews on its own <http://www.nytimes.com> **Web site**.

The decision clearly gave freelance authors electronic reproduction rights, providing they didn't have written...

... Of course, then the National Writers Union, plaintiff in the Tasini case, took out an **ad** urging writers to not take the Times up on that offer unless they got compensated...

...across the Web picking up information as they go? If The New York Times' own **Web site** has lost all its book reviews, then why not use the ones you find on...of the material.

Another Solution

Specifically, online commercial services and searchable database archives on publisher **Web sites** should continue to maintain the **inverted file index** terms and tags that identify material barred from full-text delivery by the Tasini decision. The **inverted file** indexes belong to the host services, regardless of the fact that all the terms were generated from text produced by authors, freelance or otherwise. If **inverted file** indexes remain complete and comprehensive, they could continue to identify relevant articles, at least by...

... Once the linear file is created, search engine software processes the text to generate an **inverted file index**. In the case of full-text databases, that usually means taking every word in the...

... back to the full linear file document record. When users search, they only use the **inverted file index** until they create a set of search results. When they display all or part of...

...back to the original linear file to gather the documents.

In this proposal, the underlying **inverted file index** upon which the searching process rests would continue to retain all the index terms generated...

... articles. But since no one could ever re-create a whole document from using the **inverted file** indexing, that indexing constitutes a new creation and one copyrighted to the online service. In...

... search services, usually that just means that someone has shut down the links between the **inverted file index** and the linear file containing the documents as documents. In most cases, services will only...

... complete as possible. We all know that no full-text archive is really complete-no **ads**, usually no graphics, usually no letters to the editor,

often no short news items, sometimes...comprehensive after all-for once.

Here and now, I promise all commercial hosts and publisher **Web sites** that I, for one, will beat my drums as loud as I can beat them...

... new ballgame. I call on searchers everywhere who agree with this approach to send me **e - mail messages** (bquint@mindspring.com) of support. I promise to forward them on to the relevant executives...

... editor in chief of Searcher, contributing editor for NewsBreaks, and a longtime online searcher. Her **e - mail** address is bquint@mind spring.com.

11/3,K/3 (Item 3 from file: 15)
DIALOG(R)File 15:ABI/Inform(R)
(c) 2005 ProQuest Info&Learning. All rts. reserv.

02018841 53074465
Organizing the world's information
Pemberton, Jeff
Online v24n3 PP: 41-48 May/Jun 2000
ISSN: 0146-5422 JRNL CODE: ONL
WORD COUNT: 4226

...TEXT: engine uses patentpending PageRank technology to perform an objective measurement of the importance of **Web pages**. This measurement is calculated by solving an equation of 500 million variables and more than ...

...to commercial sites, and Google SiteSearch, which adds search capability to information within a specific **site**. **ONLINE** publisher Jeff Pemberton talked with Larry Page, now Google's CEO, to find out more...

...is proportional to the number of choices you have. So if you go to a **Web page** with a lot of links, it actually takes longer to do a search there than...

...than if you saw just a static summary.
PEMBERTON: You say you store the entire **Web** ?

PAGE : Yes, we store the entire Web that we've indexed.

PEMBERTON When a query comes...

...words occurred. That part of it is a pretty typical system.

PEMBERTON: It's an **inverted file** ?

PAGE: Yes, an **inverted index** or an **inverted file**. We still use that technology. It's good technology. It's why search engines work...

...the Web and find things. Of course, that's not very practical.

We have an **inverted index**, and we do Boolean AND queries. All the terms you use in your query have...

...good.

Another part of our technology is that we're looking at not only the **Web page**, but we're looking at the pages around it in the hypertext. So if someone...

... a malicious person wants people to think he's Stanford, so he takes Stanford's **home page** and copies it onto his own server. Just looking at the text of the page...

...plenty of clues sprinkled around. One is that lots of people link to the Stanford **home page**, and the people who link to the Stanford page also

tend to have a lot...

... not just random pages. There's plenty of quality pages that link to the Stanford **home page** , and that's significant, too. So when we run a search, we look at all...

...It's stored in our index.

PEMBERTON: So that's existing in parallel with the **inverted index** ?

PAGE: Yes, it's part of the **inverted index** . As a first approximation, that kind of information is stored in our index, so when...

... thousand links out there that point to Stanford, and it's mentioned in context. The **home page** is mentioned in all those places. It's a pretty good measure that it's...s say, Yahoo! decided ONLINE is really great, so they create a link on their **home page** . Yahoo! gets half abillion or so page views per day. That would be very significant...

...thinking about it is surfing the Web at random. If I gave you a random **Web page** and you clicked on a link at random, you would get another page.. Then you...

...link at random. Just keep doing that-click, click, click. Then, after you visit every **page** on the **Web** , I'll count up how often you've visited each one. It turns out that...

...around the Web. It doesn't take into account people searching or hearing stuff [promoting **Web sites**] on the radio or whatever, but as people see that stuff, it gets reflected in same site. Will you show only the **home page** ?

PAGE : No, we'll probably return whatever ranks highest.

PEMBERTON: How has your company grown since... is sort of a de facto Internet standard. You just create a file on your **Web site** . It has a specific name, and there's a way of describing certain documents that...

...of nice for consumers. And Yahoo!, for example, just had a great quarter based on **advertising** revenue. It's a real business making a lot of money. That seems to be...

...a business standpoint, it has a lot of benefits. You don't have to have **advertising** . You're paying content creators for whatever they've created, and that's really a...

... Yes, absolutely. The example I'll give you is RedHat. They say, "I have a **Web site** . I want search over it, but I don't really know how to do it ...we can keep innovating.

Jeff Pemberton (jeffp@onlineinc.com) is Publisher of ONLINE magazine.

Comments? **Email** letters to the Editor to editor@onlineinc.com.

11/3,K/6 (Item 2 from file: 275)
DIALOG(R)File 275:Gale Group Computer DB(TM)
(c) 2005 The Gale Group. All rts. reserv.

02503916 SUPPLIER NUMBER: 74649327 (USE FORMAT 7 OR 9 FOR FULL TEXT)
The Searching Quagmire. (Internet/Web/Online Service Information)
Feldman, Susan; Liddy, Elizabeth
Searcher, 9, 5, 66
May, 2001
ISSN: 1070-4795 LANGUAGE: English RECORD TYPE: Fulltext
WORD COUNT: 7470 LINE COUNT: 00571

... user and how these design decisions can be influenced by designers,

writers, and authors of **Web pages** .

Selection is a slippery slope. Whether we choose materials by quality of data, authority, or...have chosen. Others want to keep you on their site to get additional revenue from **advertisements** that pop up every time you search or display a page.

Correcting False Assumptions About...

...of terms to search against. This is created by gathering the text of millions of **Web pages** and then creating an alphabetical list of words with ties to their locations in each **Web page** indexed, an **inverted index** . It is the **inverted index** that users search when they use a Web search engine, not the original pages. In...

...then follows links from those pages to add others. But there are over a billion **pages** on the **Web** , and to make matters worse, those pages appear and disappear with no warning. The sheer quantity of **Web pages** makes building an effective Web search engine a Sysiphean endeavor.

To add to the confusion...

...with deep rich archives of information go online, this is an increasingly common design for **Web sites** : **Pages** are stored in a database as plain text and dynamically turned into HTML only after...

...are still millions, perhaps billions, of pages. In order to keep up with the newest **pages** , **Web** search engines often take submissions from owners or designers of pages so as to know...

...list of sites that are crawled regularly. Even so, only a small percentage of the **Web pages** are crawled. It is not feasible to crawl them all. Here again the question of...

...often? A newspaper site, crawled once a week would miss 6 days of news. Normal **Web sites** , on the other hand, may change at most once a month or even once every...AND with some latitude for finding additional good matches.

Some search engines index the whole **Web page** , while others only index the first 500-1,000 words. Some use the number of...

...fuzzy AND. Full-text indexing may be added or subtracted. And, of course, many more **Web pages** will be crawled, URLs dropped, and dynamic databases added. Ranking algorithms will be subtly altered...

...long for them to go to each page and search it separately. Consequently, if a **Web page** has been moved or removed, the fact of its nonexistence may not be discovered for...

...less than technically proficient business to sue its previous Internet Service Provider because its previous **Web page** kept being retrieved after it had already moved to another ISP. The business owners were...

...hand, yield direct answers to specific questions. In the case of Yahoo!, the best-known **directory** on the **Web** , experts select **sites** on a subject and organize the listings according to a subject scheme that groups together...

...as being good navigation tools. However, directories like Yahoo! cannot possibly contain the number of **sites** that the **Web** does, nor would one want them to. The strength of a directory lies in its...academic and research sites, while others are more product-oriented.

* Factors that increase revenues from **advertising** or other sources also improve the quality of the search results. Goto.com was the...

...the "serious" players in the field, like using the yellow pages and scanning for boxed **ads** or at least ones in boldface.

However, business relationships or **advertisements** can more subtly skew the information offered. While other search engines besides Goto.com claim...

...adjoining column on the results page. This may provide links to other

related sites, to **advertising** partners who sell related products, or to collections about the topic being searched. For instance...have been giving higher rankings to sites that pay for being returned. Similarly, spamdexing by **Web sites** in order to obtain high rankings or inappropriate listings has also angered searchers. However, the...

...a high weight in search.

- * How many links there are to a page from other **pages** on the **Web** .

- * Other unpredictable factors like popularity (Direct Hit), whether the site has paid for its ranking...

...in trouble.

- * The whole Web is crawled.

- * All search engines search the same set of **Web pages** .

- * Search engines search the actual page at the time of searching.

- * All search engines work...

...to deal with it: the author of either the document or the metadata or the **Web page** design.

Web search engines dance a wary minuet ...unwitting user may see. Sometimes this is useful -- since Web crawlers index only text while **Web sites** may want to present words as graphics, e.g., corporate logos, the search engine page...

...are better matches.

- * Goto.com sells high rankings to the highest bidder.

- * Finding company's **home page** is sometimes difficult. Many search engines have tried to resolve such searches ahead of time...

...MSN (Microsoft Network), use RealNames for this purpose. RealNames promises to select the one right **home page site** for

PRODUCT/INDUSTRY NAMES: 4811525 (**Online Search Services & Directories**)

?

.?t s15/3,k/3

15/3,K/3 (Item 3 from file: 15)
DIALOG(R)File 15:ABI/Inform(R)
(c) 2005 ProQuest Info&Learning. All rts. reserv.

02029621 54926414

Inside a search engine

Anonymous

IEEE Spectrum v37n6 PP: 92-93 Jun 2000

ISSN: 0018-9235 JRNL CODE: SPC

...ABSTRACT: perform the same basic function: starting with a query, they point to the best-matching **Web pages**. To speed up the matching process search engines continuously sift through millions of **pages** on the **Web**, creating a huge central index of words and phrases against which the query is compared. To build their indexes, search engines typically use crawler. Most engines build an **inverted index**, in which each word points directly to all the documents it occurs in.

?

? ds

Set	Items	Description
S1	121779	ADVERTIS? OR ADVERTIZ? OR AD OR ADS OR PROMOTION? OR ADVERT? T? ? OR COMMERCIAL()MESSAGE?
S2	4336	(INVERTED OR FORWARD) (2N) (INDEX? ? OR INDICES OR FILE OR - FILES OR LIST OR LISTS OR STRUCTURE?)
S3	31545	(INTERNET OR WEB OR ONLINE OR ON()LINE OR HOME) (2N) (PAGE OR PAGES OR SITE OR SITES OR PORTAL? OR DIRECTOR?)
S4	57803	(E OR ELECTRONIC OR DIGITAL OR VIRTUAL) (1W) (MAIL??? OR M- ESSAG??? OR CORRESPOND?) OR EMAIL???? OR (INTERNET OR ON()LI- NE OR ONLINE OR WEB) (1W) MAIL???? OR MIME OR SMTP OR POP(1N)- MAIL
S5	69	S1(S)S2
S6	25	S5(S) (S3 OR S4)
S7	10	S6 AND IC=G06F?
S8	29	S5 (S) (INTERNET OR WEB OR WWW OR ONLINE OR ON()LINE)
S9	11	S8 AND IC=G06F?
S10	2	S9 NOT S7

? show files

File 348:EUROPEAN PATENTS 1978-2004/Dec W03

(c) 2004 European Patent Office

File 349:PCT FULLTEXT 1979-2002/UB=20041230,UT=20041223

(c) 2004 WIPO/Univentio

?

t s7/3,k/3,4

7/3,K/3 (Item 3 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2004 WIPO/Univentio. All rts. reserv.

00858320

METHOD AND APPARATUS FOR IDENTIFYING RELATED SEARCHES IN A DATABASE SEARCH SYSTEM

PROCEDE ET DISPOSITIF POUR L'IDENTIFICATION DE RECHERCHES CONNEXES DANS UN SYSTEME D'INTERROGATION DE BASE DE DONNEES

Patent Applicant/Assignee:

GOTO COM INC, 140 W. Union Street, Pasadena, CA 91103, US, US (Residence)
, US (Nationality)

Inventor(s):

ROREX Phillip G, 25115 Summerhill Lane, Stevenson Ranch, CA 91381, US,
SOULANILLE Thomas A, 931 S. El Molino Avenue, Pasadena, CA 91106, US,
HAUGAARD Bradley R, 324 Stedman Place, Monrovia, CA 91016, US,

Legal Representative:

RAUCH John G (agent), Brinks Hofer Gilson & Lione, P.O. Box 10087,
Chicago, IL 60610, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200190947 A1 20011129 (WO 0190947)

Application: WO 2001US16161 20010518 (PCT/WO US0116161)

Priority Application: US 2000575894 20000522

Designated States:

(Protection type is "patent" unless otherwise stated - for applications prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL
TJ TM TR TT TZ UA UG UZ VN YU ZA ZW

(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR

(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG

(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW

(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext Word Count: 15029

Main International Patent Class: G06F-017/30

International Patent Class: G06F-017/60

Fulltext Availability:

Detailed Description

Detailed Description

... the searches of block 400 and block 402 is a listing of rows of the **inverted index** or indexes containing the searched information. Each row contains the information associated with a search listing of the pay for performance database along with all the text of the **web page** associated with the search listing. In the illustrated embodiment, the search listing includes the **advertiser**'s search terms, the URL of the **web page**, a title and descriptive text.

At block 404, the returned related search results are sorted...set canon cnt \$cnt where

cannon-search-text = \$cst" >

</SQL>

</SQL>

```
</SQL>
</a>
</script>
Aggregates web page body-text and
listings based on the related-search
result, while collecting and creating
derived-data of 1, how many different
advertisers have web - pages associated
with the related-search result.

<script language=vortex>
<timeout = -1></timeout>
<DB = /home/goto...
```

7/3,K/4 (Item 4 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2004 WIPO/Univentio. All rts. reserv.

00824135 **Image available**

DATA STREAMING

TRANSMISSION DE DONNEES

Patent Applicant/Assignee:

BANDWIZ INC, 100 Lowder Brook Drive, Suite 1300, Westwood, MA 02090, US,
US (Residence), US (Nationality), (For all designated states except:
US)

Patent Applicant/Inventor:

RAJWAN Doron, 47 Borohov Street, 53221 Givataim, IL, IL (Residence), IL
(Nationality), (Designated only for: US)

LUBETZKY Eyal, 39 Ben Tzvi Street, 53631 Givataim, IL, IL (Residence), IL
(Nationality), (Designated only for: US)

Legal Representative:

FENSTER Paul (et al) (agent), Fenster & Company Patent Attorneys, Ltd.,
P.O. Box 10256, 49002 Petach Tikva, IL,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200157667 A1 20010809 (WO 0157667)

Application: WO 2001IL106 20010202 (PCT/WO IL0100106)

Priority Application: US 2000179926 20000203; US 2000217139 20000710; IL
137624 20000801; IL 138114 20000827; US 2000245000 20001101; US
2000245098 20001102; IL 140504 20001224

Parent Application/Grant:

Related by Continuation to: US 2000245098 20001102 (CIP)

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE DK DM DZ EE
ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT
LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM
TR TT TZ UA UG US UZ VN YU ZA ZW

(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR

(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG

(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW

(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext Word Count: 11338

Main International Patent Class: G06F-011/00

International Patent Class: G06F-013/00 ...

Fulltext Availability:

Claims

Claim

... playback function. Optionally, for points in the

1 6

file where playback is expected, the **file structure** is **inverted** in time, with earlier blocks being short and/or transmitted more often, so that playback...assist such a solution. Two particular examples of such a slowly changing channel is the **Internet WWW page** of CNN (which is widely viewed) and the CNN daily continuous newscast. In addition to... additionally, a user may activate fast forward and/or frame skipping functions (e.g., skip **advertisements**) in order that his display also catches up with the real-time event. Alternatively...

?

?ds

Set	Items	Description
S1	2574	ADVERTIS? OR ADVERTIZ? OR AD OR ADS OR PROMOTION? OR ADVERT? T? ? OR COMMERCIAL()MESSAGE?
S2	51	(INVERTED OR FORWARD OR BLOCK) (2N) (INDEX? ? OR INDICES OR FILE OR FILES OR LIST OR LISTS OR STRUCTURE?)
S3	9850	(INTERNET OR WEB OR ONLINE OR ON()LINE OR HOME) (2N) (PAGE OR PAGES OR SITE OR SITES OR PORTAL? OR DIRECTOR?)
S4	1	S1 AND S2 AND S3
S5	1	S1 AND S2
S6	2	INVERT?(2N) (INDEX? OR INDICES OR FILE OR FILES OR FILING OR LIST OR LISTS OR LISTING OR STRUCTURE?)

?show files

File 256:TecInfoSource 82-2004/Dec

(c) 2004 Info.Sources Inc

?ds

Set	Items	Description
S1	15	GOOGLE (S) (INVERTED())INDEX)
S2	11	RD (unique items)

?show files

File 2:INSPEC 1969-2004/Dec W2
(c) 2004 Institution of Electrical Engineers

File 8:Ei Compendex(R) 1970-2005/Dec W4
(c) 2005 Elsevier Eng. Info. Inc.

File 13:BAMP 2005/Dec W4
(c) 2005 The Gale Group

File 88:Gale Group Business A.R.T.S. 1976-2005/Jan 03
(c) 2005 The Gale Group

File 275:Gale Group Computer DB(TM) 1983-2005/Jan 05
(c) 2005 The Gale Group

File 349:PCT FULLTEXT 1979-2002/UB=20041230,UT=20041223
(c) 2004 WIPO/Univentio

File 545:Investext(R) 1982-2005/Jan 05
(c) 2005 Thomson Financial Networks

File 654:US Pat.Full. 1976-2005/Han 04
(c) Format only 2005 The Dialog Corp.

File 993:NewsRoom 2002
(c) 2004 The Dialog Corporation

File 995:NewsRoom 2000
(c) 2004 The Dialog Corporation

?

?t s2/3,k/1-4,6-11

2/3,K/1 (Item 1 from file: 2)

DIALOG(R)File 2:INSPEC

(c) 2004 Institution of Electrical Engineers. All rts. reserv.

7685970 INSPEC Abstract Number: C2003-08-7250R-074

Title: I/O-efficient techniques for computing Pagerank

Author(s): Chan, Y.-Y.; Gan, C.; Suel, T.

Author Affiliation: Dept. of CIS, Polytech. Univ., Brooklyn, NY, USA

Conference Title: Proceedings of the Eleventh International Conference on Information and Knowledge Management. CIKM 2002 p.549-57

Editor(s): Kalpakis, K.; Goharian, N.; Grossman, D.

Publisher: ACM, New York, NY, USA

Publication Date: 2002 Country of Publication: USA xiv+690 pp.

ISBN: 1 58113 492 4 Material Identity Number: XX-2003-00889

U.S. Copyright Clearance Center Code: 1 58113 492 4/2002/0011...\$5.00

Conference Title: ACM CIKM 2002, 11th International Conference on Information and Knowledge Management

Conference Sponsor: ACM

Conference Date: 4-9 Nov. 2002 Conference Location: Mclean, VA, USA

Language: English

Subfile: C

Copyright 2003, IEE

...Abstract: results. One widely known approach is the Pagerank technique, which forms the basis of the **Google** ranking scheme, and which assigns a global importance measure to each page based on the...

...means that it can be precomputed and then used to optimize the layout of the **inverted index** structure accordingly. However, computing the Pagerank measure requires implementing an iterative process on a massive...

2/3,K/2 (Item 1 from file: 13)

DIALOG(R)File 13:BAMP

(c) 2005 The Gale Group. All rts. reserv.

1164410 Supplier Number: 02434707 (USE FORMAT 7 OR 9 FOR FULLTEXT)

Organizing the World's Information: Google Raises the Bar on Search

Technology

(The technical and commercial success of one search engine is discussed in this interview with the company's CEO, who says a smart search engine is ideal)

Article Author(s): Pemberton, Jeff; Pack, Thomas

Online Magazine, v 24, n 3, p 41-48

May 2000

DOCUMENT TYPE: Journal; Interview & speech ISSN: 0146-5422 (United States)

LANGUAGE: English RECORD TYPE: Fulltext; Abstract

WORD COUNT: 4229

ABSTRACT:

...Google SiteSearch, which adds search capability to information within a specific site. Google uses an **inverted index** technology and searches not only the Web page, but also its hypertext links to execute...

2/3,K/3 (Item 1 from file: 88)

DIALOG(R)File 88:Gale Group Business A.R.T.S.

(c) 2005 The Gale Group. All rts. reserv.

06067825 SUPPLIER NUMBER: 83444474

Knowledge encapsulation for focused search from pervasive devices.

Aridor, Yariv; Carmel, David; Maarek, Yoelle S.; Soffer, Aya; Lempel, Ronny
ACM Transactions on Information Systems, 20, 1, 25(22)

Jan, 2002

ISSN: 1046-8188

LANGUAGE: English

RECORD TYPE: Fulltext; Abstract

WORD COUNT: 8780 LINE COUNT: 00778

... index is downloaded to the device as part of the KAB. Given a query, the **inverted index** is used to locate the KAB pages that best answer this query. The results are...

...This is similar in spirit to the PageRank (Brin and Page 1998) functionality of the **Google** search engine.) Due to the locality of the **inverted index**, ranked search results are returned to the user in a response time of less than...

2/3,K/4 (Item 1 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2004 WIPO/Univentio. All rts. reserv.

01065490 **Image available**

DATA STORE FOR KNOWLEDGE-BASED DATA MINING SYSTEM
MAGASIN DE DONNEES POUR SYSTEME D'EXPLORATION DE DONNEES BASE SUR LA
CONNAISSANCE

Patent Applicant/Assignee:

INTERNATIONAL BUSINESS MACHINES CORPORATION, New Orchard Road, Armonk, NY
10504, US, US (Residence), US (Nationality)

IBM UNITED KINGDOM LIMITED, PO Box 41, North Harbour, Portsmouth,
Hampshire PO6 3AU, GB, GB (Residence), GB (Nationality), (Designated
only for: MG)

Inventor(s):

DENESUK Matthew, 1770 Hudson Drive, San Jose, CA 95124, US,
GRUHL Daniel Frederick, 6675 Dorene Place, San Jose, CA 95120, US,
MCCURLEY Kevin Snow, 6721 Tannahill Drive, San Jose, CA 95120, US,
MEYER Joerg, 3193 Kilo Avenue, San Jose, CA 95124, US,
RAJAGOPALAN Sridhar, 19206 Vineyard Lane, Saratoga, CA 95070, US,
TOMKINS Andrew, 737 Hary Road, San Jose, CA 95120, US,
ZIEN Jason Yeong, 928 Wright Avenue No 408, Mountain View, CA 95043, US,

Legal Representative:

LITHERLAND David Peter (agent), IBM United Kingdom Limited, Intellectual
Property Law, Hursley Park, Winchester, Hampshire SO21 2JN, GB,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200396220 A1 20031120 (WO 0396220)

Application: WO 2003GB1800 20030428 (PCT/WO GB0301800)

Priority Application: US 2002142673 20020508

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NI NO NZ OM PH PL PT RO RU SC SD SE
SG SK SL TJ TM TN TR TT TZ UA UG UZ VC VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE
SI SK TR

(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG

(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW

(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext Word Count: 7228

Fulltext Availability:

Detailed Description

Detailed Description

... through to find the information he or she desires.

Existing search engines such as AltaVista, **Google**, Northern Light,
FAST, and Inktomi work by "crawling" the Web, i.e., they access Web pages
and pages to which the accessed pages hyperlink, generating an **inverted
index** of words that occur on the Web pages. The index correlates words
with the identifications...

2/3,K/6 (Item 1 from file: 654)
DIALOG(R)File 654:US Pat.Full.
(c) Format only 2005 The Dialog Corp. All rts. reserv.

5498921 **IMAGE Available
Derwent Accession: 2004-141367

Utility

E/ Information extraction from a database

Inventor: Brin, Sergey, Palo Alto, CA
Assignee: Google Inc. (02), Mountain View, CA
Google Inc
Examiner: Metjahic, Safet (Art Unit: 211)
Assistant Examiner: Al-hashemi, Sara
Law Firm: Harrity & Snyder, LLP.

	Publication Number	Kind	Date	Application Number	Filing Date
	-----	--	-----	-----	-----
Main Patent	US 6678681	A	20040113	US 2000521996	20000309

Fulltext Word Count: 7081

Description of the Invention:

...was used. This data is part of the Stanford WebBase and is used for the **GOOGLE** search engine, which includes an embodiment of the invention. As a part of the search engine, an **inverted index** of the entire repository was built...

2/3,K/7 (Item 2 from file: 654)
DIALOG(R)File 654:US Pat.Full.
(c) Format only 2005 The Dialog Corp. All rts. reserv.

5453592 **IMAGE Available
Derwent Accession: 2003-896461

Utility

E/ Detecting duplicate and near-duplicate files

Inventor: Pugh, William, Kensington, MD
Henzinger, Monika H., Menlo Park, CA
Assignee: Google, Inc. (02), Mountain View, CA
Google Inc
Examiner: Coby, Frantz (Art Unit: 211)
Assistant Examiner: Nguyen, Merilyn
Law Firm: Straub & Pokotylo
Combined Principal Attorneys: Pokotylo, John C.

	Publication Number	Kind	Date	Application Number	Filing Date
	-----	--	-----	-----	-----
Main Patent	US 6658423	A	20031202	US 2001768947	20010124

Fulltext Word Count: 12817

Description of the Invention:

...may access the stored content 154 and may generate a content index (e.g., an **inverted index**, to be described below) and content ratings (e.g., PageRanks, to be described below) 140...will be used with an advanced search facility, such as the one presently available from **Google**, Inc. of Mountain View, Calif. FIG. 2 is a process bubble diagram of such an

2/3,K/8 (Item 3 from file: 654)
DIALOG(R)File 654:US Pat.Full.
(c) Format only 2005 The Dialog Corp. All rts. reserv.

0005433140 **IMAGE Available
Derwent Accession: 2003-876244
Data store for knowledge-based data mining system

Inventor: Denesuk, Matthew, INV
Gruhl, Daniel, INV
McCurley, Kevin, INV
Meyer, Joerg, INV
Rajagopalan, Sridhar, INV
Tomkins, Andrew, INV
Zien, Jason, INV

Assignee: International Business Machines Corporation (02), Armonk, NY
Correspondence Address: John L. Rogitz Rogitz & Associates, Suite 3120 750
B Street, San Diego, CA, 92101, US

	Publication Number	Kind	Date	Application Number	Filing Date
	-----	--	-----	-----	-----
Main Patent	US 20030212699	A1	20031113	US 2002142673	20020508

Fulltext Word Count: 9778

Summary of the Invention:

0003] Existing search engines such as AltaVista, **Google**, Northern Light, FAST, and Inktomi work by "crawling" the Web, i.e., they access Web pages and pages to which the accessed pages hyperlink, generating an **inverted index** of words that occur on the Web pages. The index correlates words with the identifications...

2/3,K/9 (Item 4 from file: 654)

DIALOG(R)File 654:US Pat.Full.

(c) Format only 2005 The Dialog Corp. All rts. reserv.

0005433116 **IMAGE Available
Derwent Accession: 2003-853402
Knowledge-based data mining system

Inventor: Denesuk, Matthew, INV
Gruhl, Daniel, INV
McCurley, Kevin, INV
Rajagopalan, Sridhar, INV
Tomkins, Andrew, INV

Assignee: International Business Machines Corporation (02), Armonk, NY
Correspondence Address: John L. Rogitz Rogitz & Associates, Suite 3120 750
B Street, San Diego, CA, 92101, US

	Publication Number	Kind	Date	Application Number	Filing Date
	-----	--	-----	-----	-----
Main Patent	US 20030212675	A1	20031113	US 2002141327	20020508

Fulltext Word Count: 8962

Summary of the Invention:

0003] Existing search engines such as AltaVista, **Google**, Northern Light, FAST, and Inktomi work by "crawling" the Web, i.e., they access Web pages and pages to which the accessed pages hyperlink, generating an **inverted index** of words that occur on the Web pages. The index correlates words with the identifications...

2/3,K/10 (Item 5 from file: 654)

DIALOG(R)File 654:US Pat.Full.

(c) Format only 2005 The Dialog Corp. All rts. reserv.

0005433090 **IMAGE Available
Derwent Accession: 2003-853396

Knowledge-based data mining system

Inventor: Denesuk, Matthew, INV
Gruhl, Daniel, INV
McCurley, Kevin, INV
Rajagopalan, Sridhar, INV
Tomkins, Andrew, INV

Assignee: International Business Machines Corporation (02), Armonk, NY, US
Correspondence Address: John L. Rogitz Rogitz & Associates, Suite 3120 750
B Street, San Diego, CA, 92101, US

	Publication Number	Kind	Date	Application Number	Filing Date
Main Patent	US 20030212649	A1	20031113	US 2002141686	20020508

Fulltext Word Count: 10034

Summary of the Invention:

0003] Existing search engines such as AltaVista, **Google**, Northern Light, FAST, and Inktomi work by "crawling" the Web, i.e., they access Web pages and pages to which the accessed pages hyperlink, generating an **inverted index** of words that occur on the Web pages. The index correlates words with the identifications...

2/3,K/11 (Item 6 from file: 654)

DIALOG(R)File 654:US Pat.Full.

(c) Format only 2005 The Dialog Corp. All rts. reserv.

5349005 **IMAGE Available
Derwent Accession: 2003-754748

Utility

E/ **Detecting query-specific duplicate documents**

Inventor: Gomes, Benedict, Berkeley, CA
Smith, Benjamin Thomas, Mountain View, CA
Assignee: Google, Inc. (02), Mountain View, CA
Google Inc

Examiner: Vu, Kim (Art Unit: 212)

Assistant Examiner: Liang, Gwen

Law Firm: Straub & Pokotylo

Combined Principal Attorneys: Pokotylo, John C.

	Publication Number	Kind	Date	Application Number	Filing Date
Main Patent	US 6615209	A	20030902	US 2000684542	20001006

Fulltext Word Count: 13187

Description of the Invention:

...may access the stored content 754 and may generate a content index (e.g., an **inverted index**, to be described below) and content rankings (e.g., PageRanks, to be described below) 740...will be used with an advanced search facility, such as the one presently available from **Google**, Inc. of Mountain View, Calif. FIG. 8 is a process bubble diagram of such an...ii) translating occurrences of words in the current document into hit lists. To generate the **inverted index** 826, the indexing/sorting process 820 may then sort the partially sorted forward index 822...be run by a web server and may use a lexicon 832, together with the **inverted index** 826 and the PageRanks 828 to generate query results in response to a query. The...the search process 830' is an advanced searching facility such as the one employed by **Google**, Inc. of Mountain View, Calif. and introduced in [selection]4.1 above, the search process 830' may use a content index (an **inverted index**) 826', page ranks 828', a lexicon 832' and a repository 814' to generate the rank-ordered query results 910. In such a case, exemplary data structures of the **inverted index** 826', the page ranks 828', the lexicon 832' and

· " the repository 814' may be the...
?

'
?ds

Set	Items	Description
S1	10	AU='HARIK GEORGES':AU='HARIK GEORGES R'

?show files

File 347:JAPIO Nov 1976-2004/Aug(Updated 041203)
(c) 2004 JPO & JAPIO

File 348:EUROPEAN PATENTS 1978-2004/Dec W03
(c) 2004 European Patent Office

File 349:PCT FULLTEXT 1979-2002/UB=20041230,UT=20041223
(c) 2004 WIPO/Univentio

File 350:Derwent WPIX 1963-2004/UD,UM &UP=200482
(c) 2004 Thomson Derwent

?

?t sl/5/1-10

1/5/1 (Item 1 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2004 European Patent Office. All rts. reserv.

01744949

METHOD AND APPARATUS FOR CHARACTERIZING DOCUMENTS BASED ON CLUSTERS OF RELATED WORDS

PROCEDE ET DISPOSITIF DESTINES A CARACTERISER DES DOCUMENTS SUR LA BASE DE GROUPES DE MOTS ASSOCIES

PATENT ASSIGNEE:

Google Inc., (4836400), 2400 Bayshore Parkway, Mountain View, CA 94043,
(US), (Applicant designated States: all)

INVENTOR:

HARIK, Georges , 900 High School Way 2307, Mountain View, CA 94041,
(US)

SHAZEER, Noam, M., 245 Easy St. Apt. 1, Mountain View, CA 94043, (US
PATENT (CC, No, Kind, Date):

WO 2004031916 040415

APPLICATION (CC, No, Date): EP 2003774573 031003; WO 2003US31545 031003

PRIORITY (CC, No, Date): US 416144 P 021003

DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR;
HU; IE; IT; LI; LU; MC; NL; PT; RO; SE; SI; SK; TR

EXTENDED DESIGNATED STATES: AL; LT; LV; MK

INTERNATIONAL PATENT CLASS: G06F-001/00

LEGAL STATUS (Type, Pub Date, Kind, Text):

Application: 040609 A2 International application. (Art. 158(1))

Application: 040609 A2 International application entering European
phase

LANGUAGE (Publication,Procedural,Application): English; English; English

1/5/2 (Item 2 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2004 European Patent Office. All rts. reserv.

01742384

SERVING ADVERTISEMENTS BASED ON CONTENT

PRESENTATION D'ANNONCES BASEE SUR LE CONTENU

PATENT ASSIGNEE:

Google Inc., (4836400), 2400 Bayshore Parkway, Mountain View, CA 94043,
(US), (Applicant designated States: all)

INVENTOR:

ANDERSON, Darrell, 111 N. Rengstorff Ave., Apt. 167, Mountain View, CA
94043, (US)

BUCHHEIT, Paul, 900 High School Way, Apt. 2324, Mountain View, CA 94041,
(US)

CAROBUS, Alexander Paul, 900 High School Way, Apt. 2 137, Mountain View,
CA 94041, (US)

CUI, Yingwei, 6092 Willowgrove Lane, Cupertino, CA 95014, (US)

DEAN, Jeffrey A., 884 Fifteenth Avenue, Menlo Park, CA 94025, (US)

HARIK, Georges R. , 95 High School Way, Apt. 3135, Mountain View, CA
94041, (US)

JINDAL, Deepak, 151 Calderon Avenue, Apt. 291, Cupertino, CA 94041, (US)

SHIVAKUMAR, Narayanan, 1093 McKay Drive, Menlo Park, CA 95131, (US)

PATENT (CC, No, Kind, Date):

WO 2004028234 040408

APPLICATION (CC, No, Date): EP 2003798730 030924; WO 2003US30233 030924

PRIORITY (CC, No, Date): US 413536 P 020924; US 314427 021206; US 375900
030226

DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR;
HU; IE; IT; LI; LU; MC; NL; PT; RO; SE; SI; SK; TR

EXTENDED DESIGNATED STATES: AL; LT; LV; MK

INTERNATIONAL PATENT CLASS: G06F-001/00

LEGAL STATUS (Type, Pub Date, Kind, Text):

Application: 040602 A2 International application. (Art. 158(1))

Application: 040602 A2 International application entering European

phase
LANGUAGE (Publication,Procedural,Application): English; English; English

1/5/3 (Item 3 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2004 European Patent Office. All rts. reserv.

01741458
SERVING ADVERTISEMENTS USING INFORMATION ASSOCIATED WITH E-MAIL
FOURNITURE D'ANNONCES PUBLICITAIRES AU MOYEN D'INFORMATIONS ASSOCIEES A UN
MESSAGE ELECTRONIQUE
PATENT ASSIGNEE:
Google Inc., (4836400), 2400 Bayshore Parkway, Mountain View, CA 94043,
(US), (Applicant designated States: all)
INVENTOR:
DEAN, Jeffrey, A., 884 Fifteenth Avenue, Menlo Park, CA 94025, (US)
Harik, Georges R. , 950 High School Way, Apt. 3135, Mountain View, CA
94041, (US)
Bucheit, Paul, 900 High School Way, Apt. 2324, Mountain View, CA 94041,
(US)
PATENT (CC, No, Kind, Date):
WO 2004029759 040408
APPLICATION (CC, No, Date): EP 2003759524 030924; WO 2003US30235 030924
PRIORITY (CC, No, Date): US 413536 P 020924; US 314427 021206; US 375900
030226; US 452830 030602
DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR;
HU; IE; IT; LI; LU; MC; NL; PT; RO; SE; SI; SK; TR
EXTENDED DESIGNATED STATES: AL; LT; LV; MK
INTERNATIONAL PATENT CLASS: G06F-001/00
LEGAL STATUS (Type, Pub Date, Kind, Text):
Application: 040602 A2 International application. (Art. 158(1))
Application: 040602 A2 International application entering European
phase
LANGUAGE (Publication,Procedural,Application): English; English; English

1/5/4 (Item 4 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2004 European Patent Office. All rts. reserv.

01741457
METHODS AND APPARATUS FOR SERVING RELEVANT ADVERTISEMENTS
PROCEDE ET APPAREIL PERMETTANT DE SERVIR DES ANNONCES PUBLICITAIRES
PERTINENTS
PATENT ASSIGNEE:
Google Inc., (4836400), 2400 Bayshore Parkway, Mountain View, CA 94043,
(US), (Applicant designated States: all)
INVENTOR:
DEAN, Jeffrey, A., 884 Fifteenth Avenue, Menlo Park, CA 94025, (US)
HARIK, Georges, R. , 950 High School Way, Apt. 3135, Mountain View, CA
94041, (US)
BUCHEIT, Paul, 900 High School Way, Apt. 2324, Mountain View, CA 94041,
(US)
PATENT (CC, No, Kind, Date):
WO 2004029758 040408
APPLICATION (CC, No, Date): EP 2003759523 030924; WO 2003US30232 030924
PRIORITY (CC, No, Date): US 413536 P 020924; US 314427 021206
DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR;
HU; IE; IT; LI; LU; MC; NL; PT; RO; SE; SI; SK; TR
EXTENDED DESIGNATED STATES: AL; LT; LV; MK
INTERNATIONAL PATENT CLASS: G06F-001/00
LEGAL STATUS (Type, Pub Date, Kind, Text):
Application: 040602 A2 International application. (Art. 158(1))
Application: 040602 A2 International application entering European
phase
LANGUAGE (Publication,Procedural,Application): English; English; English

1/5/5 (Item 5 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2004 European Patent Office. All rts. reserv.

01741292

**SUGGESTING AND/OR PROVIDING AD SERVING CONSTRAINT INFORMATION
SUGGESTION ET/OU FOURNITURE D'INFORMATIONS DE CONTRAINTE DE SERVICE
D'ANNONCES PUBLICITAIRES**

PATENT ASSIGNEE:

Google Inc., (4836400), 2400 Bayshore Parkway, Mountain View, CA 94043,
(US), (Applicant designated States: all)

INVENTOR:

DEAN, Jeffrey, A., 884 Fifteenth Avenue, Menlo Park, CA 94025, (US)
HARIK, Georges, R., 950 High School Way, Apt. No. 3135, Mountain View,
CA 94041, (US)
BUCHEIT, Paul, 900 High School Way, Apt. No. 2324, Mountain View, CA
94041, (US)

PATENT (CC, No, Kind, Date):

WO 2004029827 040408

APPLICATION (CC, No, Date): EP 2003754885 030924; WO 2003US30234 030924

PRIORITY (CC, No, Date): US 413536 P 020924; US 314427 021206; US 389688
030314

DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR;
HU; IE; IT; LI; LU; MC; NL; PT; RO; SE; SI; SK; TR

EXTENDED DESIGNATED STATES: AL; LT; LV; MK

INTERNATIONAL PATENT CLASS: G06F-017/21; G06F-017/30; G06F-017/60

LEGAL STATUS (Type, Pub Date, Kind, Text):

Application: 040602 A1 International application. (Art. 158(1))
Application: 040602 A1 International application entering European
phase

LANGUAGE (Publication,Procedural,Application): English; English; English

1/5/6 (Item 1 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2004 WIPO/Univentio. All rts. reserv.

01110236 **Image available**

**METHOD AND APPARATUS FOR CHARACTERIZING DOCUMENTS BASED ON CLUSTERS OF
RELATED WORDS**

**PROCEDE ET DISPOSITIF DESTINES A CARACTERISER DES DOCUMENTS SUR LA BASE DE
GROUPES DE MOTS ASSOCIES**

Patent Applicant/Assignee:

GOOGLE INC, 2400 Bayshore Parkway, Mountain View, CA 94043, US, US
(Residence), US (Nationality)

Inventor(s):

HARIK Georges, 900 High School Way #2307, Mountain View, CA 94041, US,
SHAZEER Noam M, 245 Easy St. Apt. 1, Mountain View, CA 94043, US

Legal Representative:

PARK Richard (agent), 508 Second St., Ste. 201, Davis, CA 95616, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200431916 A2-A3 20040415 (WO 0431916)
Application: WO 2003US31545 20031003 (PCT/WO US03031545)
Priority Application: US 2002416144 20021003

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK
LR LS LT LU LV MA MD MG MK MN MW MX MZ NI NO NZ OM PG PH PL PT RO RU SC
SD SE SG SK SL SY TJ TM TN TR TT TZ UA UG UZ VC VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE
SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Main International Patent Class: G06F-017/30

Publication Language: English
Filing Language: English
Fulltext Availability:
Detailed Description
Claims
Fulltext Word Count: 23087

English Abstract

One embodiment of the present invention provides a system characterizes a document with respect to clusters of conceptually related words. Upon receiving a document containing a set of words, the system selects "candidate clusters" of conceptually related words that are related to the set of words (Figure 22, 2202). These candidate clusters are selected using a model that explains how sets of words are generated from clusters of conceptually related words (Figure 22, 2204). Next, the system constructs a set of components to characterize the document, wherein the set of components includes components for candidate clusters (Figure 22, 2206). Each component in the set of components indicates a degree to which a corresponding candidate cluster is related to the set of words (Figure 22, 2208).

French Abstract

Un mode de realisation de la presente invention concerne un systeme destine a caracteriser un document par rapport a des groupes de mots associes conceptuellement. Lors de la reception d'un document contenant un ensemble de mots, le systeme selectionne des "groupes candidats" de mots associes conceptuellement associes a un ensemble de mots. Ces groupes candidats sont selectionnes au moyen d'un modele expliquant comment les ensembles de mots sont generes a partir de groupes de mots associes conceptuellement. Ensuite, le systeme construit un ensemble de composants en vue de caracteriser le document, cet ensemble de composants comprenant des composants destines a des groupes candidats. Chaque composant dans l'ensemble de composants indique un degre selon lequel un groupe candidat correspondant est associe a l'ensemble de mots.

Legal Status (Type, Date, Text)

Publication 20040415 A2 Without international search report and to be republished upon receipt of that report.
Search Rpt 20041223 Late publication of international search report
Republication 20041223 A3 With international search report.
Republication 20041223 A3 Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

1/5/7 (Item 2 from file: 349)

DIALOG(R)File 349:PCT FULLTEXT

(c) 2004 WIPO/Univentio. All rts. reserv.

01107906 **Image available**

**SERVING ADVERTISEMENTS USING INFORMATION ASSOCIATED WITH E-MAIL
FOURNITURE D'ANNONCES PUBLICITAIRES AU MOYEN D'INFORMATIONS ASSOCIEES A UN
MESSAGE ELECTRONIQUE**

Patent Applicant/Assignee:

GOOGLE INC, 2400 Bayshore Parkway, Mountain View, CA 94043, US, US
(Residence), US (Nationality)

Inventor(s):

DEAN Jeffrey A, 884 Fifteenth Avenue, Menlo Park, CA 94025, US,
HARIK Georges R, 950 High School Way, Apt. #3135, Mountain View, CA
94041, US,
BUCHEIT Paul, 900 High School Way, Apt. #2324, Mountain View, CA 94041,
US

Legal Representative:

SUN Karl (agent), Google, Inc., 2400 Bayshore Parkway, Mountain View, CA
94043, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200429759 A2 20040408 (WO 0429759)
Application: WO 2003US30235 20030924 (PCT/WO US03030235)
Priority Application: US 2002413536 20020924; US 2002314427 20021206; US

2003375900 20030226; US 2003452830 20030602

Designated States:

(Protection type is "patent" unless otherwise stated - for applications prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SK
SL TJ TM TN TR TT TZ UA UG UZ VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE
SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Main International Patent Class: G06F

Publication Language: English

Filing Language: English

Fulltext Availability:

Detailed Description

Claims

Fulltext Word Count: 11194

English Abstract

Advertisers are permitted to put targeted ads on e-mails. The present invention may do so by (i) obtaining information of an e-mail that includes available spots for ads, (ii) determining one or more ads relevant to the e-mail information, and/or (iii) providing the one or more ads for rendering in association with the e-mail.

French Abstract

Les annonceurs peuvent inserer des annonces publicitaires cibles dans des messages electroniques. Pour ce faire, la presente invention consiste (i) a obtenir des informations d'un message electronique comprenant des espaces disponibles pour des annonces publicitaires, (ii) a determiner une ou plusieurs annonces publicitaires pertinentes selon les informations de message electronique, et/ou (iii) a fournir ces annonces publicitaires en vue d'une restitution en association avec ce message electronique.

Legal Status (Type, Date, Text)

Publication 20040408 A2 Without international search report and to be republished upon receipt of that report.

1/5/8 (Item 3 from file: 349)

DIALOG(R) File 349:PCT FULLTEXT

(c) 2004 WIPO/Univentio. All rts. reserv.

01107905 **Image available**

SUGGESTING AND/OR PROVIDING AD SERVING CONSTRAINT INFORMATION

SUGGESTION ET/OU FOURNITURE D'INFORMATIONS DE CONTRAINTE DE SERVICE D'ANNONCES PUBLICITAIRES

Patent Applicant/Assignee:

GOOGLE INC, 2400 Bayshore Parkway, Mountain View, CA 94043, US, US
(Residence), US (Nationality)

Inventor(s):

DEAN Jeffrey A, 884 Fifteeth Avenue, Menlo Park, CA 94025, US,
HARIK Georges R, 950 High School Way, Apt. #3135, Mountain View, CA
94041, US,
BUCHHEIT Paul, 900 High School Way, Apt. #2324, Mountain View, CA 94041,
US

Legal Representative:

SUN Karl (agent), Google, Inc., 2400 Bayshore Parkway, Mountain View, CA
94043, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200429827 A1 20040408 (WO 0429827)

Application: WO 2003US30234 20030924 (PCT/WO US03030234)

Priority Application: US 2002413536 20020924; US 2002314427 20021206; US
2003389688 20030314

Designated States:

(Protection type is "patent" unless otherwise stated - for applications prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SK
SL TJ TM TN TR TT TZ UA UG UZ VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE
SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Main International Patent Class: G06F-017/21

International Patent Class: G06F-017/30; G06F-017/60

Publication Language: English

Filing Language: English

Fulltext Availability:

Detailed Description

Claims

Fulltext Word Count: 7243

English Abstract

Targeting information (also referred to as ad "serving constraints") or candidate targeting information for an advertisement is identified (1). Targeting information may be identified (410) by extracting topics or concepts (420) from, and/or generating topics or concepts based on, ad information, such as information from a Web page to which an ad is linked (or some other Web page of interest to the ad or advertiser). (400). The topics or concepts may be relevant queries associated with the Web page of interest, clusters, etc.

French Abstract

Selon l'invention, des informations de ciblage (egalement appelees "contraintes de service" d'annonces publicitaires) et des informations de ciblage candidates pour une annonce publicitaire sont identifiees. Ces informations de ciblage peuvent etre identifiees par extraction de sujets ou de concepts a partir d'informations d'annonces publicitaires, et/ou par generation de sujets ou de concepts sur la base d'informations d'annonces publicitaires, telles que des informations issues d'une page Web a laquelle une annonce est liee (ou une autre page Web d'interet pour l'annonce ou l'annonceur). Ces sujets ou ces concepts peuvent etre des demandes pertinentes associees a la page Web d'interet, des groupes, etc.

Legal Status (Type, Date, Text)

Publication 20040408 A1 With international search report.

Publication 20040408 A1 Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

1/5/9 (Item 4 from file: 349)

DIALOG(R)File 349:PCT FULLTEXT

(c) 2004 WIPO/Univentio. All rts. reserv.

01107904 **Image available**

SERVING ADVERTISEMENTS BASED ON CONTENT

PRESENTATION D'ANNONCES BASEE SUR LE CONTENU

Patent Applicant/Assignee:

GOOGLE INC, 2400 Bayshore Parkway, Mountain View, CA 94043, US, US
(Residence), US (Nationality)

Inventor(s):

ANDERSON Darrell, 111 N. Rengstorff Ave., Apt. #167, Mountain View, CA 94043, US,
BUCHHEIT Paul, 900 High School Way, Apt. #2324, Mountain View, CA 94041, US,
CAROBUS Alexander Paul, 900 High School Way, Apt. 2#137, Mountain View, CA 94041, US,
CUI Yingwei, 6092 Willowgrove Lane, Cupertino, CA 95014, US,

DEAN Jeffrey A, 884 Fifteenth Avenue, Menlo Park, CA 94025, US,
HARIK Georges R , 95 High School Way, Apt. #3135, Mountain View, CA
94041, US,
JINDAL Deepak, 151 Calderon Avenue, Apt. #291, Cupertino, CA 94041, US,
SHIVAKUMAR Narayanan, 1093 McKay Drive, Menlo Park, CA 95131, US
Legal Representative:
SUN Karl (agent), Google, Inc., 2400 Bayshore Parkway, Mountain View, CA
94043, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200428234 A2-A3 20040408 (WO 0428234)
Application: WO 2003US30233 20030924 (PCT/WO US03030233)
Priority Application: US 2002413536 20020924; US 2002314427 20021206; US
2003375900 20030226

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SK
SL TJ TM TN TR TT TZ UA UG UZ VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE
SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Main International Patent Class: G06F-017/30

International Patent Class: G06F-017/60

Publication Language: English

Filing Language: English

Fulltext Availability:

Detailed Description

Claims

Fulltext Word Count: 14155

English Abstract

Advertisers (110) are permitted to put targeted ads on page on the web
(or some other document of any media type) (Figure 1). The present
invention may do so by (i) obtaining content that includes available
spots for ads (120), (ii) determining ads relevant to content, and/or
(iii) combining content with ads determined to be relevant to the
content.

French Abstract

Des annonceurs peuvent mettre des annonces cibles sur une page web (ou
sur d'autres documents de n'importe quel type de moyen de communication).
La presente invention concerne un procede consistant : 1) a obtenir un
contenu comprenant des messages publicitaires disponibles pour des
annonces, 2) a determiner quelles sont les annonces pertinentes par
rapport au contenu, et/ou 3) a combiner le contenu avec des annonces
considerees pertinentes par rapport au contenu.

Legal Status (Type, Date, Text)

Publication 20040408 A2 Without international search report and to be
republished upon receipt of that report.

Search Rpt 20040527 Late publication of international search report

Republication 20040527 A3 With international search report.

1/5/10 (Item 5 from file: 349)

DIALOG(R)File 349:PCT FULLTEXT

(c) 2004 WIPO/Univentio. All rts. reserv.

01107903 **Image available**

METHODS AND APPARATUS FOR SERVING RELEVANT ADVERTISEMENTS

**PROCEDE ET APPAREIL PERMETTANT DE SERVIR DES ANNONCES PUBLICITAIRES
PERTINENTS**

Patent Applicant/Assignee:

GOOGLE INC, 2400 Bayshore Parkway, Mountain View, CA 94043, US, US
(Residence), US (Nationality)

Inventor(s):

DEAN Jeffrey A, 884 Fifteenth Avenue, Menlo Park, CA 94025, US,
HARIK Georges R , 950 High School Way, Apt. #3135, Mountain View, CA
94041, US,
BUCHHEIT Paul, 900 High School Way, Apt. #2324, Mountain View, CA 94041,
US

Legal Representative:

SUN Karl (agent), Google, Inc., 2400 Bayshore Parkway, Mountain View, CA
94043, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200429758 A2-A3 20040408 (WO 0429758)
Application: WO 2003US30232 20030924 (PCT/WO US03030232)
Priority Application: US 2002413536 20020924; US 2002314427 20021206

Designated States:

(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)

AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SK
SL TJ TM TN TR TT TZ UA UG UZ VN YU ZA ZM ZW
(EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE
SI SK TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM

Main International Patent Class: G06F-017/30

Publication Language: English

Filing Language: English

Fulltext Availability:

Detailed Description

Claims

Fulltext Word Count: 6511

English Abstract

The relevance of advertisements to a user's interests is improved (Figure 4). In one implementation, the content of a web page is analyzed to determine a list of one or more topics associated with that web page (420). An advertisement is considered to be relevant (440) to that web page if it is associated with keywords belonging to the list of one or more topics. One or more of these relevant advertisements may be provided for rendering in conjunction with the web page or related web pages.

French Abstract

Selon l'invention, la pertinence des annonces publicitaires par rapport aux interets de l'utilisateur est amelioree. Dans un mode de realisation, le contenu d'une page web est analyse afin de determiner une liste d'un ou de plusieurs sujets associes a cette page web. Une annonce publicitaire est consideree pertinente pour cette page web si elle est associee a des mots cles appartenant a la liste d'un ou de plusieurs sujets. Une ou plusieurs de ces annonces peuvent etre fournies avec ladite page web ou les pages web associees pour la presentation.

Legal Status (Type, Date, Text)

Publication 20040408 A2 Without international search report and to be
republished upon receipt of that report.

Search Rpt 20040527 Late publication of international search report

Republication 20040527 A3 With international search report.

Republication 20040527 A3 Before the expiration of the time limit for
amending the claims and to be republished in the
event of the receipt of amendments.

?

?ds

Set	Items	Description
S1	1	AU='HARIK, GEORGES RAIF'
S2	5	AU='HARIK, G.R.':AU='HARIK, G.R.'
S3	6	S1:S2

?show files

File 35:Dissertation Abs Online 1861-2004/Dec
(c) 2004 ProQuest Info&Learning

File 65:Inside Conferences 1993-2005/Jan W1
(c) 2005 BLDSC all rts. reserv.

File 148:Gale Group Trade & Industry DB 1976-2005/Jan 05
(c)2005 The Gale Group

File 2:INSPEC 1969-2004/Dec W2
(c) 2004 Institution of Electrical Engineers

File 16:Gale Group PROMT(R) 1990-2005/Jan 05
(c) 2005 The Gale Group

File 160:Gale Group PROMT(R) 1972-1989
(c) 1999 The Gale Group

File 636:Gale Group Newsletter DB(TM) 1987-2005/Jan 05
(c) 2005 The Gale Group

File 624:McGraw-Hill Publications 1985-2004/Dec 28
(c) 2004 McGraw-Hill Co. Inc

File 275:Gale Group Computer DB(TM) 1983-2005/Jan 05
(c) 2005 The Gale Group

File 647:CMP Computer Fulltext 1988-2005/Dec W3
(c) 2005 CMP Media, LLC

File 674:Computer News Fulltext 1989-2004/Dec W2
(c) 2004 IDG Communications

File 9:Business & Industry(R) Jul/1994-2005/Jan 04
(c) 2005 The Gale Group

File 610:Business Wire 1999-2005/Jan 05
(c) 2005 Business Wire.

File 810:Business Wire 1986-1999/Feb 28
(c) 1999 Business Wire

?

?t s3/6,/1-6

3/6/1 (Item 1 from file: 35)

01578064 ORDER NO: AAD97-32090

**LEARNING GENE LINKAGE TO EFFICIENTLY SOLVE PROBLEMS OF BOUNDED DIFFICULTY
USING GENETIC ALGORITHMS (PROBLEM SOLVING, STOCHASTIC OPTIMIZATION)**

Year: 1997

3/6/2 (Item 1 from file: 2)

6983696 INSPEC Abstract Number: C2001-08-1180-092

Title: Compressed introns in a linkage learning genetic algorithm

Publication Date: 1998

Copyright 2001, IEE

3/6/3 (Item 2 from file: 2)

6631900 INSPEC Abstract Number: C2000-08-1180-026

Title: Linkage learning through probabilistic expression

Publication Date: 9 June 2000

Copyright 2000, IEE

3/6/4 (Item 3 from file: 2)

6504427 INSPEC Abstract Number: C2000-03-1180-091

Title: A parameter-less genetic algorithm

Publication Date: 1999

Copyright 2000, IEE

3/6/5 (Item 4 from file: 2)

6411650 INSPEC Abstract Number: B2000-01-0260-014, C2000-01-1180-014

Title: The compact genetic algorithm

Publication Date: Nov. 1999

Copyright 1999, IEE

3/6/6 (Item 5 from file: 2)

6016610 INSPEC Abstract Number: B9810-0260-035, C9810-1180-046

Title: The compact genetic algorithm

Publication Date: 1998

Copyright 1998, IEE

?